

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
ptnr:SP TREMBL- ACC:Q9D398	6330415E02RIK PROTEIN - Mus musculus (Mouse)	945	862/945 (91%)	897/945 (94%)	0.0
ptnr:SP TREMBL- ACC:O08722	TRANSMEMBRANE RECEPTOR UNC5H2	945	862/945 (91%)	893/945 (94%)	0.0
ptnr:SP TREMBL- ACC:O08747	UNC-5 HOMOLOG (C. ELEGANS)	931	610/929 (65%)	723/929 (77%)	0.0
ptnr:SP TREMBL- ACC:O95185	TRANSMEMBRANE RECEPTOR UNC5C - Homo sapiens	931	598/929 (64%)	718/929 (77%)	0.0

5

10

- 15

25

35

40

45

5	NOV1b	ENSKWSACSTECAHWRSRECMAPPPGGRDCSGTLLDSNNCTDGLCNC-----TB-SGEALYAGLVVAF	374
	Q9D398	ENSKWSACSTECAHWRSRECMAPPPGGRDCSGTLLDSNNCTDGLVNLQRTLNPKSPHLESCVVALYAGLVVAF	386
	008722	ENSKWSACSTECAHWRSRECMAPPPGGRDCSGTLLDSNNCTDGLVNLQRTLNPKSPHLESCVVALYAGLVVAF	386
	008747	ENSKWSACSTECAHWRSRECMAPPPGGRDCSGTLLDSNNCTDGLCNC-----APDSGEALYAGLVVAF	389
095185	ENSKWSACSTECAHWRSRECMAPPPGGRDCSGTLLDSNNCTDGLCNC-----APDSGEALYAGLVVAF	389	
10	NOV1a	WAILLMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIRGCPVYALQCSAOKI	454
	NOV1b	WAILLMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIRGCPVYALQCSAOKI	454
	Q9D398	WAILLMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIRGCPVYALQCSAOKI	466
	008722	WAILLMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIRGCPVYALQCSAOKI	466
008747	WAILLMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIRGCPVYALQCSAOKI	463	
095185	WAILLMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIRGCPVYALQCSAOKI	463	
15	NOV1a	PMNTNSPLLDPLPSLKVYVSSSTCGSGCLADGADLLGVLPCTYPSDFRHTFHLHRSASLGSOQLLGLPRDFSSVS	534
	NOV1b	PMNTNSPLLDPLPSLKVYVSSSTCGSGCLADGADLLGVLPCTYPSDFRHTFHLHRSASLGSOQLLGLPRDFSSVS	534
	Q9D398	PMNTNSPLLDPLPSLKVYVSSSTCGSGCLADGADLLGVLPCTYPSDFRHTFHLHRSASLGSOQLLGLPRDFSSVS	546
	008722	PMNTNSPLLDPLPSLKVYVSSSTCGSGCLADGADLLGVLPCTYPSDFRHTFHLHRSASLGSOQLLGLPRDFSSVS	546
008747	PMNTNSPLLDPLPSLKVYVSSSTCGSGCLADGADLLGVLPCTYPSDFRHTFHLHRSASLGSOQLLGLPRDFSSVS	546	
095185	PMNTNSPLLDPLPSLKVYVSSSTCGSGCLADGADLLGVLPCTYPSDFRHTFHLHRSASLGSOQLLGLPRDFSSVS	533	
20	NOV1a	GTFGCLGGRLLPTGCGVSLVLPAGAPGKGFYEMYLINNAESTPLPSEGTQVLSPSVTCGPTGLLLCRPVVLTPHCA	614
	NOV1b	GTFGCLGGRLLPTGCGVSLVLPAGAPGKGFYEMYLINNAESTPLPSEGTQVLSPSVTCGPTGLLLCRPVVLTPHCA	614
	Q9D398	GTFGCLGGRLLPTGCGVSLVLPAGAPGKGFYEMYLINNAESTPLPSEGTQVLSPSVTCGPTGLLLCRPVVLTPHCA	626
	008722	GTFGCLGGRLLPTGCGVSLVLPAGAPGKGFYEMYLINNAESTPLPSEGTQVLSPSVTCGPTGLLLCRPVVLTPHCA	626
008747	GTFGCLGGRLLPTGCGVSLVLPAGAPGKGFYEMYLINNAESTPLPSEGTQVLSPSVTCGPTGLLLCRPVVLTPHCA	613	
095185	GTFGCLGGRLLPTGCGVSLVLPAGAPGKGFYEMYLINNAESTPLPSEGTQVLSPSVTCGPTGLLLCRPVVLTPHCA	613	
25	NOV1a	EVSARQWIFOLKTOAHQGHWEVVTDEETLTPCYCOLEPRACHILLDLQGTYYFTGESYSRSRAVKRLQIAFPALCIT	694
	NOV1b	EVSARQWIFOLKTOAHQGHWEVVTDEETLTPCYCOLEPRACHILLDLQGTYYFTGESYSRSRAVKRLQIAFPALCIT	694
	Q9D398	EVSARQWIFOLKTOAHQGHWEVVTDEETLTPCYCOLEPRACHILLDLQGTYYFTGESYSRSRAVKRLQIAFPALCIT	706
	008722	EVSARQWIFOLKTOAHQGHWEVVTDEETLTPCYCOLEPRACHILLDLQGTYYFTGESYSRSRAVKRLQIAFPALCIT	706
008747	EVSARQWIFOLKTOAHQGHWEVVTDEETLTPCYCOLEPRACHILLDLQGTYYFTGESYSRSRAVKRLQIAFPALCIT	693	
095185	EVSARQWIFOLKTOAHQGHWEVVTDEETLTPCYCOLEPRACHILLDLQGTYYFTGESYSRSRAVKRLQIAFPALCIT	693	
30	NOV1a	SLSEYLRVYCLEDTFVALKEVLELERTLGGYVLEEPNLFYFKDSYHNLRLSLHDTLPHAWRSKLLAKYQEIFFYHWSGSE	774
	NOV1b	SLSEYLRVYCLEDTFVALKEVLELERTLGGYVLEEPNLFYFKDSYHNLRLSLHDTLPHAWRSKLLAKYQEIFFYHWSGSE	774
	Q9D398	SLSEYLRVYCLEDTFVALKEVLELERTLGGYVLEEPNLFYFKDSYHNLRLSLHDTLPHAWRSKLLAKYQEIFFYHWSGSE	786
	008722	SLSEYLRVYCLEDTFVALKEVLELERTLGGYVLEEPNLFYFKDSYHNLRLSLHDTLPHAWRSKLLAKYQEIFFYHWSGSE	786
008747	SLSEYLRVYCLEDTFVALKEVLELERTLGGYVLEEPNLFYFKDSYHNLRLSLHDTLPHAWRSKLLAKYQEIFFYHWSGSE	773	
095185	SLSEYLRVYCLEDTFVALKEVLELERTLGGYVLEEPNLFYFKDSYHNLRLSLHDTLPHAWRSKLLAKYQEIFFYHWSGSE	773	
35	NOV1a	QNALHCTFTLERHSLASTELTCNIVCRQVEGEGQIFOLHTTLAETPAGSLDLCASAPGNTTOLGPYAFKIPLSITROKI	854
	NOV1b	QNALHCTFTLERHSLASTELTCNIVCRQVEGEGQIFOLHTTLAETPAGSLDLCASAPGNTTOLGPYAFKIPLSITROKI	854
	Q9D398	QNALHCTFTLERHSLASTELTCNIVCRQVEGEGQIFOLHTTLAETPAGSLDLCASAPGNTTOLGPYAFKIPLSITROKI	866
	008722	QNALHCTFTLERHSLASTELTCNIVCRQVEGEGQIFOLHTTLAETPAGSLDLCASAPGNTTOLGPYAFKIPLSITROKI	852
008747	QNALHCTFTLERHSLASTELTCNIVCRQVEGEGQIFOLHTTLAETPAGSLDLCASAPGNTTOLGPYAFKIPLSITROKI	852	
095185	QNALHCTFTLERHSLASTELTCNIVCRQVEGEGQIFOLHTTLAETPAGSLDLCASAPGNTTOLGPYAFKIPLSITROKI	852	
40	NOV1a	CSSLDAPIRGCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEANQDDGDLASLASALEENGKSEMLVATDGGC	933
	NOV1b	CSSLDAPIRGCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEANQDDGDLASLASALEENGKSEMLVATDGGC	945
	Q9D398	CSSLDAPIRGCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEANQDDGDLASLASALEENGKSEMLVATDGGC	945
	008722	CSSLDAPIRGCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEANQDDGDLASLASALEENGKSEMLVATDGGC	945
008747	CSSLDAPIRGCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEANQDDGDLASLASALEENGKSEMLVATDGGC	931	
095185	CSSLDAPIRGCHDWRMLAOKLSMDRYLNYFATKASPTGVILDLWEANQDDGDLASLASALEENGKSEMLVATDGGC	931	

The presence of identifiable domains in NOV1, as well as all other NOVX proteins, was determined by searches using software algorithms such as PROSITE, DOMAIN, Blocks, Pfam, ProDomain, and Prints, and then determining the Interpro number by crossing the domain match (or numbers) using the Interpro website (<http://www.ebi.ac.uk/interpro>). DOMAIN results for NOV1 as disclosed in Tables 1G-1O, were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST analyses. This BLAST analysis software samples domains found in the Smart and Pfam collections. For Tables 1G-1O and all successive DOMAIN sequence alignments, fully conserved single residues are indicated by black shading or by the sign (!) and "strong" semi-conserved residues are indicated by grey shading or by the sign (+). The "strong" group of conserved amino acid residues may be any one of the following groups of amino acids: STA, NEQK, NHQK, NDEO, OHRK, MILV, MILF, HY, FYW.

Precursor (EC 3.1.3.48)) (E = 0.0). Public amino acid databases include the GenBank databases, SwissProt, PDB and PIR.

NOV2c is expressed in at least Synovium/Synovial membrane, Kidney. Expression information was derived from the tissue sources of the sequences that were included in the derivation of the sequence of CuraGen Acc. No. CG50718-05. The sequence is predicted to be expressed in the *Rattus norvegicus* :glomerular mesangial. because of the expression pattern of (GENBANK-ID: gb:GENBANK-ID:AF063249|acc:AF063249.1) a closely related *Rattus norvegicus* glomerular mesangial cell receptor protein-tyrosine phosphatase precursor (PTPRQ) mRNA, complete cds homolog.

Homologies among each of the above NOV2 proteins will be shared by the other NOV2 proteins insofar as they are homologous to each other as shown below in Table 2G. Any reference to NOV2 is assumed to refer to all three of the NOV2 proteins in general, unless otherwise noted.

Table 2G Alignment of NOV2a, b, and c

		10	20	30	40	50	60	
NOV2a		MDFLIIIFLLLF	IGTSETQVDVSMVVP	TRYDITISSISITTYT	SPVTRIGASN	EPGPPV	58	
NOV2b		MDFLIIIFLLLF	IGTSETQVDVSMVVP	TRYDITISSISITTYT	SPVTRIGASN	EPGPPV	1	
NOV2c		MDFLIIIFLLLF	IGTSETQVDVSMVVP	TRYDITISSISITTYT	SPVTRIVTINVT	EPGPPV	60	
		70	80	90	100	110	120	
NOV2a		FLAGERVGSAGILL	SWNTPPNPNGRII	SYIVKYKEVCPWM	QTVYTQVRSKPD	SLEVLLTN	118	
NOV2b		FLAGERVGSAGILL	SWNTPPNPNGRII	SYIVKYKEVCPWM	QTVYTQVRSKPD	SLEVLLTN	1	
NOV2c		FLAGERVGSAGILL	SWNTPPNPNGRII	SYIVKYKEVCPWM	QTVYTQVRSKPD	SLEVLLTN	120	
		130	140	150	160	170	180	
NOV2a		LNPGTTYEIKVAA	ENSAGIGVFSDF	PFLFOTAES	APGKVVDFTGE	AVPFSS	KLMWYTS	175
NOV2b		LNPGTTYEIKVAA	ENSAGIGVFSDF	PFLFOTAES	APGKVVDFTGE	AVPFSS	KLMWYTS	1
NOV2c		LNPGTTYEIKVAA	ENSAGIGVFSDF	PFLFOTAES	APGKVVDFTGE	AVPFSS	KLMWYTS	180
		190	200	210	220	230	240	
NOV2a		ATKKKITSFKISV	KHNRSGIIVKES	SIRVECHLSAS	LPFLHCNENSES	FLWSTASPSPTL	C	235
NOV2b		ATKKKITSFKISV	KHNRSGIIVKES	SIRVECHLSAS	LPFLHCNENSES	FLWSTASPSPTL	C	1
NOV2c		QPNGKITSFKISV	KHNRSGIIVKES	SIRVEDILTCKL	PECNVENSES	FLWSTASPSPTL	C	240
		250	260	270	280	290	300	
NOV2a		RVTPPSRTTHSS	STLTONEISSV	KEPISFVVTHLR	PYTYLFEVSAAT	TEAGYIDSTIV		294
NOV2b		RVTPPSRTTHSS	STLTONEISSV	KEPISFVVTHLR	PYTYLFEVSAAT	TEAGYIDSTIV		1
NOV2c		RVTPPSRTTHSS	STLTONEISSV	KEPISFVVTHLR	PYTYLFEVSAAT	TEAGYIDSTIV		300
		310	320	330	340	350	360	
NOV2a		RTPEVPEGPPQNC	VTCNITGKSFS	ILWDPPTIVTGK	FSYRVELYGPS	SGRILDNSTKDL		354
NOV2b		RTPEVPEGPPQNC	VTCNITGKSFS	ILWDPPTIVTGK	FSYRVELYGPS	SGRILDNSTKDL		1
NOV2c		RTPEVPEGPPQNC	VTCNITGKSFS	ILWDPPTIVTGK	FSYRVELYGP	SGRILDNSTKDL		359
		370	380	390	400	410	420	
NOV2a		KFAFTNLTPFTMY	DVYIAAETSAGT	GPKSNISVFTPP	DPVGAFFDLQLA	EVSTOVRITW		414

	NOV2b	1
	NOV2c	KFAFTNLTPFTMYDVYIAAETSAGTGPKSNISVFTPPDVP GAVFDLQLAEVESTQVRITW	419
5		430 440 450 460 470 480	
	NOV2a	KKPRQPNGI INQYRVKVLVPETGII LENTLLTGNN EINDPMAPEIVNIVCPMVGLYEGSA	474
	NOV2b	1
	NOV2c	KKPRQPNGI INQYRVKVLVPETGII LENTLLTGNN EINDPMAPEIVNIVCPMVGLYEGSA	479
10		490 500 510 520 530 540	
	NOV2a	EMSSDLHSLATFIYN SHPKNF PARNRAEDQTS PVVTTRNQYITDIAAEOLTYVRI RLRR	534
	NOV2b	1
	NOV2c	EMSSDLHSLATFIYN SHPKNF PARNRAEDQTS PVVTTRNQYITDIAAEOLSYVRI RLVP	539
15		550 560 570 580 590 600	
	NOV2a	FWAETMGFSRYTIMSSASRDNITSPG---ELSAQNP RVTIVTIDIEVFLWDPPP---PVF	588
	NOV2b	1
20	NOV2c	FTEHMTSVSAFTIMGEGPPTVLSVRTRQQVPSIRIINYKNTSSSLLYWDPPPYPNGK	599
		610 620 630 640 650 660	
25	NOV2a	FHHYLTITLDVENQSKSII LRTINSLSTVL IGLKKYTKYKMRVAASTHVGESSLSEENDI	648
	NOV2b	1
	NOV2c	ITHYLTIVAFELDINRAFOITITIDNSFLATG IGLKKYTKYKMRVAASTHVGESSLSEENDI	659
		670 680 690 700 710 720	
30	NOV2a	FVRTSEDEPESSPQDVEVIDVTADEIRLKWSPPEKPNGII IAYEVLYKNIDTLYMKNTST	708
	NOV2b	1
	NOV2c	FVRTSEDEPESSPQDVEVIDVTADEIRLKWSPPEKPNGII IAYEVLYKNIDTLYMKNTST	719
35		730 740 750 760 770 780	
	NOV2a	TDIILRLNLRPHTLYNISVRSYTRFGHGNQVSSLLSVRTSEIVPDSAPENITYKNISSGEI	768
	NOV2b	1
	NOV2c	TDIILRLNLRPHTLYNISVRSYTRFGHGNQVSSLLSVRTSEIVPDSAPENITYKNISSGEI	779
40		790 800 810 820 830 840	
	NOV2a	ELSFLPPSSPNGII IQYTIY LKRSNGNEERTINTTSLTQNIKGLKKYTQYII EVSASTLK	828
	NOV2b	1
45	NOV2c	ELSFLPPSSPNGII IQYTIY LKRSNGNEERTINTTSLTQNI--LKKYTQYII EVSASTLK	837
		850 860 870 880 890 900	
50	NOV2a	GEGVRSAPISILTEEDAPDSPQDFSVKQLSGVTVKLSWQPPLEPNGIILYYTVYVWR--	886
	NOV2b	1
	NOV2c	GEGVRSAPISILTEEDAPDSPQDFSVKQLSGVTVKLSWQPPLEPNGIILYYTVYVWRNR	897
		910 920 930 940 950 960	
55	NOV2a	SSLKTINVTETSL ELSDL DYNVEYSAYVTASTRFGDGKTRSNIISFOTPEGSPDPPKDVY	946
	NOV2b	1
	NOV2c	SSLKTINVTETSL ELSDL DYNVEYSAYVTASTRFGDGKTRSNIISFOTPEGSPDPPKDVY	957
60		970 980 990 1000 1010 1020	
	NOV2a	YANLSSSSII LFWTPPSKPNGII IQYYSVYYRNTSGTFMQNFTLHEVTNDFDNMTVSTIID	1006
	NOV2b	1
	NOV2c	YANLSSSSII LFWTPPSKPNGII IQYYSVYYRNTSGTFMQNFTLHEVTNDFDNMTVSTIID	1017
65		1030 1040 1050 1060 1070 1080	
	NOV2a	KLTIFSYITFWLTASTSVGNGNKSSDII EVYTDQDVPEGFVGNLTYESISSTAINVSWVP	1066
	NOV2b	26
	NOV2c	KLTIFSYITFWLTASTSVGNGNKSSDII EVYTDQDVPEGFVGNLTYESISSTAINVSWVP	1077
70		1090 1100 1110 1120 1130 1140	

5
10
15
20
25
30
35
40
45
50
55
60
65
70

NOV2a
NOV2b
NOV2c

PAQPNGLVFYVSLILQQTPRHVRPPLVTYERSIYFDNLEKYTDYILKITPSTEGFSDT 1126
PAQPNGLVFYVSLILQQTPRHVRPPLVTYERSIYFDNLEKYTDYILKITPSTEGFSDT 86
PAQPNGLVFYVSLILQQTPRHVRPPLVTYERSIYFDNLEKYTDYILKITPSTEGFSDT 1137

1150 1160 1170 1180 1190 1200

NOV2a
NOV2b
NOV2c

YTAQLYIKTEEDHPETSPIINTFKNLSSTSVLLSWDPPVKPNGAII SYDLTLQGPENYS 1186
YTAQLYIKTEEDVPETSPIINTFKNLSSTSVLLSWDPPVKPNGAII SYDLTLQGPENYS 146
YTAQLYIKTEEDVPETSPIINTFKNLSSTSVLLSWDPPVKPNGAII SYDLTLQGPENYS 1197

1210 1220 1230 1240 1250 1260

NOV2a
NOV2b
NOV2c

FITSDNYIILEELSPFTLYSFFAAARTRKGLGPSSILFFYTDESVP LAPPQNLT LINCTS 1246
FITSDNYIILEELSPFTLYSFFAAARTRKGLGPSSILFFYTDESVP LAPPQNLT LINCTS 206
FITSDNYIILEELSPFTLYSFFAAARTRKGLGPSSILFFYTDESVP LAPPQNLT LINCTS 1257

1270 1280 1290 1300 1310 1320

NOV2a
NOV2b
NOV2c

DFVWLKWSPLPGGIVKVYSFKIHEHETDTIYYKNISGFKTEAKLVGLEPVS TY SIRVS 1306
DFVWLKWSPLPGGIVKVYSFKIHEHETDTIYYKNISGFKTEAKLVGLEPVS TY SIRVS 266
DFVWLKWSPLPGGIVKVYSFKIHEHETDTIYYKNISGFKTEAKLVGLEPVS TY SIRVS 1317

1330 1340 1350 1360 1370 1380

NOV2a
NOV2b
NOV2c

AFTKVGNQNSNVVKFTTQESVPDVQNMQCMATSWQSVLVKWDPPKKANGII TOYMT 1366
AFTKVGNQNSNVVKFTTQESVPDVQNMQCMATSWQSVLVKWDPPKKANGII TOYMT 326
AFTKVGNQNSNVVKFTTQESVPDVQNMQCMATSWQSVLVKWDPPKKANGII TOYMT 1377

1390 1400 1410 1420 1430 1440

NOV2a
NOV2b
NOV2c

VERNSTKVSPQDHMYTFIKLLANTSYVFKVRASTASAGEDESTCHVSTLPETVPSVPTNI 1426
VERNSTKVSPQDHMYTFIKLLANTSYVFKVRASTASAGEDESTCHVSTLPETVPSVPTNI 386
VERNSTKVSPQDHMYTFIKLLANTSYVFKVRASTASAGEDESTCHVSTLPETVPSVPTNI 1437

1450 1460 1470 1480 1490 1500

NOV2a
NOV2b
NOV2c

AFSDVQSTSATLTWIRPDTILGYFQNYKITTLRAQCKEWESEECVEYQKIQVLYEAHL 1486
AFSDVQSTSATLTWIRPDTILGYFQNYKITTLRAQCKEWESEECVEYQKIQVLYEAHL 446
AFSDVQSTSATLTWIRPDTILGYFQNYKITTLRAQCKEWESEECVEYQKIQVLYEAHL 1497

1510 1520 1530 1540 1550 1560

NOV2a
NOV2b
NOV2c

TEETVYGLKKFRWYRFQVAASTNAGYGNASNWISTKTLPGPPDGPPENVHVATSPFISIS 1546
TEETVYGLKKFRWYRFQVAASTNAGYGNASNWISTKTLPGPPDGPPENVHVATSPFISIS 506
TEETVYGLKKFRWYRFQVAASTNAGYGNASNWISTKTLPGPPDGPPENVHVATSPFISIS 1557

1570 1580 1590 1600 1610 1620

NOV2a
NOV2b
NOV2c

ISWSEPAVITGPTCYLIDVKSVDNDEFNISFIKSNEENKTIEIKDLEIFTRYSVVITAF 1606
ISWSEPAVITGPTCYLIDVKSVDNDEFNISFIKSNEENKTIEIKDLEIFTRYSVVITAF 566
ISWSEPAVITGPTCYLIDVKSVDNDEFNISFIKSNEENKTIEIKDLEIFTRYSVVITAF 1617

1630 1640 1650 1660 1670 1680

NOV2a
NOV2b
NOV2c

GNISAAVVEGKSSAEMIVTTLESAPKDPNNMTFOKIPDEVTKFQLTSLPPSQPNGNIQV 1666
GNISAAVVEGKSSAEMIVTTLESAPKDPNNMTFOKIPDEVTKFQLTSLPPSQPNGNIQV 626
GNISAAVVEGKSSAEMIVTTLESAPKDPNNMTFOKIPDEVTKFQLTSLPPSQPNGNIQV 1677

1690 1700 1710 1720 1730 1740

NOV2a
NOV2b
NOV2c

YQALVYREDDPTAVQIHNLISIIQKTNTFVIAMLEGLKGGHTYINISVYAVNSAGAGPKVPM 1726
YQALVYREDDPTAVQIHNLISIIQKTNTFVIAMLEGLKGGHTYINISVYAVNSAGAGPKVPM 686
YQALVYREDDPTAVQIHNLISIIQKTNTFVIAMLEGLKGGHTYINISVYAVNSAGAGPKVPM 1737

1750 1760 1770 1780 1790 1800

NOV2a
NOV2b
NOV2c

RITMDIKAPARPKTKPTPIYDATGKLLVTSTTITIRMPICYSDDHGPIKNVQVLVTETG 1786
RITMDIKAPARPKTKPTPIYDATGKLLVTSTTITIRMPICYSDDHGPIKNVQVLVTETG 746
RITMDIKAPARPKTKPTPIYDATGKLLVTSTTITIRMPICYSDDHGPIKNVQVLVTETG 1797

		1810	1820	1830	1840	1850	1860	
5	NOV2a	1846
	NOV2b	806
	NOV2c	1857
		1870	1880	1890	1900	1910	1920	
10	NOV2a	1906
	NOV2b	855
	NOV2c	1917
		1930	1940	1950	1960	1970	1980	
15	NOV2a	1962
	NOV2b	855
	NOV2c	1977
		1990	2000	2010	2020	2030	2040	
20	NOV2a	2017
	NOV2b	855
	NOV2c	2037
25		2050	2060	2070	2080	2090	2100	
	NOV2a	2077
	NOV2b	855
	NOV2c	2097
		2110	2120	2130	2140	2150	2160	
35	NOV2a	2137
	NOV2b	855
	NOV2c	2157
		2170	2180	2190	2200	2210	2220	
40	NOV2a	2197
	NOV2b	855
	NOV2c	2217
		2230	2240	2250	2260	2270	2280	
45	NOV2a	2257
	NOV2b	855
	NOV2c	2277
50		2290	2300					
	NOV2a					2281 (SEQ ID NO:6)
	NOV2b					855 (SEQ ID NO:8)
55	NOV2c					2300 (SEQ ID NO:10)

The disclosed NOV2a polypeptide has homology to the amino acid sequences shown in the BLASTP data listed in Table 2H.

NOV2A-MDFLLIPLLLFIGSETQVDSNVVPGRYDIITSSIS--TTVTSPTF
gi|12621078| MMDFFHFSFLPLIIGTSESQVDSVSSFDGIGYDILSSVSA-TTSSSPVSH
gi|125977|-MGLQMTAARPTAALLLVLSLLLTWTHPTIVDAAPPEIIRK
gi|10728878|-MGLQMTAARPTAALLLVLSLLLTWTHPTIVDAAPPEIIRK
gi|7290546|-MDCATRKQQQLRAHHKQQQIQIQTHGRKKRQQQKQRKHNNHHYVONSOQQ
gi|1362625|-MDCATRKQQQLRAHHKQQQIQIQTHGRKKRQQQKQRKHNNHHYVONPOQQ

60 70 80 90 100

NOV2AIGAS--NEPGPPVFLAGERVGSAGILLISWNTPPNPNCRIISYIYKYKEVC
gi|12621078| TLATNVTKPPPPVFLAGERVGSAGILLISWNTPPNPNCRIISYVVKYKEVC
gi|125977| PONGVRVGGVASFYCAARGDPPPSVWRKNGKKVSG
gi|10728878| PONGVRVGGVASFYCAARGDPPPSVWRKNGKKVSG
gi|7290546| QKFVWLNVGCLLIHLAQHANAADLVI--NVPNASSNANAFYRIDVSPPF
gi|1362625| QKFVWLNVGCLLIHLARHANAADLVI--NVPNASSNANAFYRIDVSPPF

110 120 130 140 150

NOV2APIMOTVYTVQVRKPDLSLEVLITLNIPTCTIYEIKVAAENSAGIGVFSDFPL
gi|12621078| PIMOTAVTRARAKPDLSLEVLITLNIPTCTIYEIKVAAENAGIGVFSDFPL
gi|125977| -T-OSRYTVLEQPGGISILRIEPPVACRDADPYECVAENGVCDAVSADAT
gi|10728878| -T-OSRYTVLEQPGGISILRIEPPVACRDADPYECVAENGVCDAVSADAT
gi|7290546| GPPENPTTIPASDIE-KDINKFSRALPGTEVNFWDYYTNSGTHREQLTWTVN
gi|1362625| GPPENPTTIPASDIE-KDINKFSRALPGTEVNFWDYYTNSGTHREQLTWTVN

160 170 180 190 200

NOV2AFCTAESAPGKVVPTFGCAVPFSSK-LMWYTS-ATKKYITSEFKISVKHNRS
gi|12621078| FCTAESAPGKVVNLIVFALNYSAVNLIWYLRQPNGKITSEFKISVKHARS
gi|125977| LTIYE-----GDKTPAGFPVITCGPTRVIEVGHITVLMTCKAIG
gi|10728878| LTIYEGWQ---KTAISGDKTIPAGFPVITCGPTRVIEVGHITVLMTCKAIG
gi|7290546| ITTAPDPP---ANLSVQLRSKSAFTIWRPP--GSGRYGGERIIVMLGLTD
gi|1362625| ITTAPDPP---ANLSVQLRSKSAFTIWRPP--GSGRYGGERIIVMLGLTD

210 220 230 240 250

NOV2AGIVVREVSIRVECISSASLPLHCNENSESEFLWSTASPSPTLGRVTPPSRT
gi|12621078| GIVVQVSLRVEDIISGKLPECNENSESEFLWSTISPSPTLGRVTPTVST
gi|125977| NPTPNYIWKVQTKVDMSENRYSKDCFLQENREEDCKYECVAENSM
gi|10728878| NPTPNYIWKVQTKVDMSENRYSKDCFLQENREEDCKYECVAENSM
gi|7290546| LPFERSYSLEGNETIOLSAK--ELTFGGSYQOVQYYSVYOCKESVAYTSN
gi|1362625| LPFERSYSLEGNETIOLSAK--ELTFGGSYQOVQYYSVYOCKESVAYTSN

260 270 280 290 300

NOV2ATHSSSTLTQNETISSV-KEPISFVYTHLRPYTTYLFEVSAATTEACYIDST
gi|12621078| TQSSSTAARSKISSVWKEPISFVYTHLRPYTTYLFEVSAVTEACYIDST
gi|125977| CTEHSKATNLVYKVRVPPTIFSRPPEITISEVMCGSNLNLSCIAVGSMPMH
gi|10728878| CTEHSKATNLVYKVRVPPTIFSRPPEITISEVMCGSNLNLSCIAVGSMPMH
gi|7290546| FTIKPNTPGKFIIVWRNETILLVLWQPPFPAGVITHYRVSITPDDATOSV
gi|1362625| FTIKPNTPGKFIIVWRNETILLVLWQPPFPAGVITHYRVSITPDDATOSV

310 320 330 340 350

NOV2AIVRTFESVPGPPFONCVTEMTTKSSESLIWDPPITVTKGFSYRVELYGPS
gi|12621078| IVRTFESVPGPPFONCIMGNTCKAESISWDPPITVTKGFSYRVELYGP
gi|125977| VKWVGSEDLTPENEMPIGRNVLOLINIQESAN
gi|10728878| VKWVGSEDLTPENEMPIGRNVLOLINIQESAN
gi|7290546| IYVEREGEPPGPAOAAFKGLPGREYNISVQT
gi|1362625| IYVEREGEPPGPAOAAFKGLPGREYNISVQT

360 370 380 390 400

NOV2AAGRILDNSTKDLKFAFTNLTPTFTMYDVVYAAETSAAGTGPKSNISVFTFPD
gi|12621078| SGRILDNSTKDLRFATHTLTPTFTMYDVVYAAETSAAGVGPKSNISVFTFPD
gi|125977|YLCIAASTLCQIDSVSVVKVGS
gi|10728878|YLCIAASTLCQIDSVSVVKVGS
gi|7290546| VSEDENS-SVPTIDARYITVPER
gi|1362625| VSEDENS-SVPTIDARYITVPER


```

      410      420      430      440      450
NOV2A  VEGAVFDILOAEVESTQVRIITWKKPRQNGIINQYRVVLPETGIITEN
gi|12621078| VEGAVFDILOAEVESTQVRIITWKKPRQNGIISQYRVVSVLEETGVVTEN
5 gi|125977| LFTAPTDOVISEVTATSVRIEWSYKG-----PEDLYYVVICYKPKQAN
gi|10728878| LFTAPTDOVISEVTATSVRIEWSYKG-----PEDLYYVVICYKPKQAN
gi|7290546| VLNVTDEAYITSSG--FRVRNEPFR--TYSEFDAYQVVLSTSRITFN
gi|1362625| VLNVTDEAYITSSG--FRVRNEPFR--TYSEFDAYQVVLSTSRITFN

      460      470      480      490      500
NOV2A  TLLTGNNR-INDPMAPEIVNIVQEMVGLYEGSAEMSSDLHSLATFIYNH
gi|12621078| TLLTGQDEISINPMSPEIMNLVDEMGIFVEGSGEMSSDLHSPASFIYNH
15 gi|125977| QAFSEISG---IITMYVVRALSEYTEYEFYVIAVNNIGRG-----
gi|10728878| QAFSEISG---IITMYVVRALSEYTEYEFYVIAVNNIGRG-----
15 gi|7290546| VPRANGF---SVYFDYDILDEGRTYEVVVKTLADNVN-----
gi|1362625| VPRANGF---SVYFDYDILDEGRTYEVVVKTLADNVN-----

      510      520      530      540      550
NOV2A  PDKNFPAARNRAEDQTSPPVITRNQYITDIAAEQLTYVLIRLRRFWAETMG
gi|12621078| PHNDFPASTRAEEOSSPVVITRNQYMTDITAEQLSYVVRRLVPFTEHTIS
gi|125977| -----PPSAPATCTIGETKMESAP-----
25 gi|10728878| -----PPSAPATCTIGETEMESAP-----
gi|7290546| ---SWPASGEVTLRPRPVRSLG-----
gi|1362625| ---SWPASGEVTLRPRPVRSLG-----

      560      570      580      590      600
NOV2A  FSRYTIMSSASRDNLTPG---PLSAQNFVRVTHVTIMEVFHMDPPFVP
30 gi|12621078| VSAFTIMGEGPPTVLTVRTREQVPSSIQINYNMISSSSLLYMDPPEYP
gi|125977| -----RNVQVTLSSSTVITWEPPETP-----
gi|10728878| -----RNVQVTLSSSTVITWEPPETP-----
35 gi|7290546| -----GFLDDG---SNALHISWEPAETG
gi|1362625| -----GFLDDG---SNALHISWEPAETG

      610      620      630      640      650
NOV2A  FFHHYLIITLDVENOSKS---IILRTLNSLSLVHIGLKKYTKYKRVAA
40 gi|12621078| NGKITHYDIYATELDINR---AFQMTVDNSFLITGLKKYTRYKRVAA
gi|125977| NGQVTGYKVYTTINSNQPEASWNSQMVNDSELTTVSDVTPHAIYTVRVOA
gi|10728878| NGQVTGYKVYTTINSNQPEASWNSQMVNDSELTTVSELTPHAIYTVRVOA
45 gi|7290546| RQDSYRISYHEQTNASEV---PAPFPVAAESQITTNLTETLDSLAGR
gi|1362625| RQDSYRISYHEQTNASEV---PAPFPVAAESQITTNLTETLDSLAGR

      660      670      680      690      700
NOV2A  STNYGESSLSEENDIFVRTSEDEPESSPDVEVIDVTADEIRLKWSPPEK
50 gi|12621078| STNYGESSLSEENDIFVRTPEDEPESSPDQVQTGVSPSELRLKWSPPEK
gi|125977| YTSMCAGPMSTP-----
gi|10728878| YTSMCAGPMSTP-----
55 gi|7290546| RYLHAVORLSKG-----
gi|1362625| RYLHAVORLSKG-----

      710      720      730      740      750
NOV2A  PNGIIIAEVLYKNIDTLYMKNTSTTDIILRNLRPHLYNISVRSYTFPG
60 gi|12621078| PNGIIIAEVLYQNADTLFVKNTSTTDIISDLPKPYTLYNISIRSITRLG
gi|125977| -----VOYKAQQGVF-----
gi|10728878| -----VOYKAQQGVF-----
65 gi|7290546| -----VASNASDIT-----R-YTRP-
gi|1362625| -----VASNASDIT-----R-YTRP-

      760      770      780      790      800
NOV2A  HGNQVSSLLSVRTSESVPDSAPENITYKNISSCEHELSLPPSSPNGIIO
70 gi|12621078| HGNQSSLLSVRTSETVPLSAPENITYKNISSCEHELSLPPSPNGIIO
gi|125977| -----SQPSNFRATDGETATLQATKPTHSSENIIV
gi|10728878| -----SQPSNFRATDGETATLQATKPTHSSENIIV
gi|7290546| -----AAPLIOELASIDEG---LMLFWRSQVNSRQD

```



```

gi|1362625| -----AAPLIQELRSTDCG---DMLSWRSDVNSRQD
                        810      820      830      840      850
5  NOV2A      KYTFLKRSNG-NEBRTTNTTSLTONIKGLKKYTYVITVSASLXGEGY
gi|12621078| KYTFLKRSNS-HEARTTNTTSLTOTIGLKKYTHYVTEVSASLXGEGY
gi|125977|    HYELVNDYVANOAHHKRISNSEAYTIDGLYEDILVYIWLAAARSORGECA
gi|10728878| HYELVNDYVANOAHHKRISNSEAYTIDGLYEDILVYIWLAAARSORGECA
gi|7290546|   RYEVHYCRNGT-REBRTNATNETSLTIHYLHEGSGYEVKRVHAISH---GV
10 gi|1362625| RYEVHYCRNGT-REBRTNATNETSLTIHYLHEGSGYEVKRVHAISH---GV
                        860      870      880      890      900
15 NOV2A      RSRPISILTEEDAPDSPPQDFSVKOLSGVTVMLSWOPP--LEPNCIILYY
gi|12621078| RSRPISILTEEDAPDSPPQDFSVKOLSGVTVMLSWOPP--LEPNCIILYY
gi|125977|    TUPPIPVRTKQYVPCAPPRNITAIATSTTISLSWLPPPVERSNCRITVY
gi|10728878| TUPPIPVRTKQYVPCAPPRNITAIATSTTISLSWLPPPVERSNCRITVY
gi|7290546|   RSEPHSYFOAVFP--KPPONLTLCVHNLVVLHWQAP--EGST-FSEYV
20 gi|1362625| RSEPHSYFOAVFP--KPPONLTLCVHNLVVLHWQAP--EGST-FSEYV
                        910      920      930      940      950
25 NOV2A      TVYVWR---SSLKITN-VTETSLELSDLDYNVEYSAYVTASTREFGCKT
gi|12621078| TVYVWR---SSLKITN-VTETSLELSDLDYNVEYSAYVTASTREFGCKT
gi|125977|    KVFVVEVGREDDEATTMTLNMTSTVLDELKRWTEYKTVVLACTSVGDC-P
gi|10728878| KVFVVEVGREDDEATTMTLNMTSTVLDELKRWTEYKTVVLACTSVGDC-P
gi|7290546|   VRYRTDA---SPWQRISGLHENEARKDKHYGERYLVCQNTVS-FGVESP
30 gi|1362625| VRYRTDA---SPWQRISGLHENEARKDKHYGERYLVCQNTVS-FGVESP
                        960      970      980      990      1000
35 NOV2A      RSNITISFCTPEG-PSDPPKDVYYANLSSSSIILFWTPPS--KPNGIIOYY
gi|12621078| RSNITISFCTPEG-PSDPPKDVYYANLSSSSIILFWTPPS--KPNGIIOYY
gi|125977|    RSHPIILRTCEDVPCD-PQDVKATPLNSTSIHVSWKPPLEKDRNGIIRGY
gi|10728878| RSHPIILRTCEDVPCD-PQDVKATPLNSTSIHVSWKPPLEKDRNGIIRGY
gi|7290546|   HPLELNVTMEPQ---PVSNNVPLVDSRNLILEWP-----RDCCHVDFY
40 gi|1362625| HPLELNVTMEPQ---PVSNNVPLVDSRNLILEWP-----RDCCHVDFY
                        1010     1020     1030     1040     1050
45 NOV2A      SVYYRNTSGTFMQNFTLHEVTNDPDNMTVSTIIDKLTFISYYTFWLTAST
gi|12621078| SVYYRNTSGTFMQNFTLHEVTNDPDNMTVSTIIDKLTFISYYTFWLTAST
gi|125977|    -----
gi|10728878| -----
50 gi|7290546| -----
gi|1362625| -----
                        1060     1070     1080     1090     1100
55 NOV2A      SVGNKNGKSSDIIIEVYTDQDVEGPGVGNLTYESISSTAINMSWVPPAPNG
gi|12621078| SVGNKNGKSSDIIIEVYTDQDVEGPGVGNLTYESISSTAINMSWVPPAPNG
gi|125977|    -----HIHAQELRDEC
gi|10728878| -----HIHAQELRDEC
gi|7290546|   -----LKKWFTDEEDR
60 gi|1362625| -----LKKWFTDEEDR
                        1110     1120     1130     1140     1150
65 NOV2A      EVFYYSGLILQOHP-RHVRPPLVTYERSHYFKNLEKYTDVILKTPSTER
gi|12621078| EVFYYSGLNLQOSP-PPRHMIPPLVTYENSIDFDELEKYTDVIFKTPSTER
gi|125977|    KGFLNEPFKFDVVD-----RUEFNVTCLOPDTKYSISVAALTRK
gi|10728878| KGFLNEPFKFDVVD-----RUEFNVTCLOPDTKYSISVAALTRK
gi|7290546|   KEFKNVITQLEDLSS-----P-SVREPTEDLSGGRGYRFEVOASSN-
70 gi|1362625| KEFKNVITQLEDLSS-----P-SVREPTEDLSGGRGYRFEVOASSN-
                        1160     1170     1180     1190     1200
75 NOV2A      GFSITYTAQHYIKKEEDIPETSPINTFENLSSFSLLSMDPVPKPNCAI
gi|12621078| GFSITYTTOHHKKEEDVDPDTPPIINTFENLSSFSLLSMDPVPKPNCAI
gi|125977|    GDCRSAAIIVKTPG-GVVPVRPTVSKIMEREPIVSDLEAERACTYCEI

```

39

gi|12621078| WISTOTLPGPDPGPPENHVAATSPFGLINISWSSEPAVITGTFITFLDVKS
gi|125977| KVTIVRIK---PEDVPLNIRAHVDVSTHSMTLWSVPPIRLT-PVNYKISFDA
gi|10728878| KVTIVRIK---PEDVPLNIRAHVDVSTHSMTLWSVPPIRLT-PVNYKISFDA
gi|7290546| TDFADVSSTLMRSSAPISASVQTLTAPPKVDYFQPSD-
gi|1362625| -----TDFADVSSTLMRSSAPISASVQTLTAPPKVDYFQPSD-

1610 1620 1630 1640 1650

NOV2A VDNDEFNISFIKSNENKTIETIKDLEIFTRYSVVITAFITGNISAAVVEGK
gi|12621078| VDDDDFNISFLKSNENKTTTEINNLEVFTRYSVVITAFVGNVSRAYTDGK
gi|125977| MK-----
gi|10728878| MK-----
gi|7290546| VQ-----
gi|1362625| VQ-----

1660 1670 1680 1690 1700

NOV2A SSAEMIVTTLESAPKDPNNNMIFOKIPDEVTKFQLTFLPPSQPNENIRVY
gi|12621078| SSAEVIITTTLESVPKDPNNNMIFOKIPDEVTKFQLTFLPPSQPNENIRVY
gi|125977| -----VFVDSQGESOT-----QIVPKREIILKH
gi|10728878| -----VFVDSQGESOT-----QIVPKREIILKH
gi|7290546| -----PGEVTFEWS-----LEPAEQHGPIIDYE
gi|1362625| -----PGEVTFEWS-----LEPAEQHGPIIDYE

1710 1720 1730 1740 1750

NOV2A QALVMREDDPTAVQIHNLSTIQKTNTFVIAMLEGKGGCHYNISVYVANS
gi|12621078| QALVMREDDPTAVQIHNLSTIQKTNTSIAMLEGKGGCHYNISVYVANS
gi|125977| YMKTHITINELSPETTYNVNVSIPS-----DYSYRPETKITVITQMAAPQ
gi|10728878| YMKTHITINELSPETTYNVNVSIPS-----DYSYRPETKITVITQMAAPQ
gi|7290546| RITCONADDADVSSYEFPVNATQG-----KIDGLVPGNHVIFRCAKSA
gi|1362625| RITCONADDADVSSYEFPVNATQG-----KIDGLVPGNHVIFRCAKSA

1760 1770 1780 1790 1800

NOV2A ACAGPKVPMRITMDIKAPARPKTKPTITDYATGKLLVTSATITIRMPICY
gi|12621078| ACAGPKVPMRITMDIKAPARPKSKPIITRDATGKLLVTSATITIRMPICY
gi|125977| PMVKPDPFYGVNGEETLVILPOASEEYGPISH-----YYLVVMPEDKSN
gi|10728878| PMVKPDPFYGVNGEETLVILPOASEEYGPISH-----YYLVVMPEDKSN
gi|7290546| LQYGAEREHLQTMPIILAPPVPEPSVTPLEVS-----TSSTIEISFRQGY
gi|1362625| LQYGAEREHLQTMPIILAPPVPEPSVTPLEVS-----TSSTIEISFRQGY

1810 1820 1830 1840 1850

NOV2A YSDDHGPPKRNQVQVIVTETGAQHDG--NVTIKYDAYENKAR-PYFTNECEP
gi|12621078| YNDDHGPPKRNQVQVIVTETGAQQDG--NVTIKYDAYENKAR-PYFTNECEP
gi|125977| LHKIPDQFLITDDLPGRNKPERPN-----APYIAAKPEPRTSEFTFLGSG
gi|10728878| LHKIPDQFLITDDLPGRNKPERPN-----APYIAAKPEPRTSEFTFLGSG
gi|7290546| FSNAGHVRYSYTIITLAEIVGKIASGLEMPSDODVCAYTVWLIPYQAIPEPN
gi|1362625| FSNAGHVRYSYTIITLAEIVGKIASGLEMPSDODVCAYTVWLIPYQAIPEPN

1860 1870 1880 1890 1900

NOV2A NPPTCEGKTKFSGNEEIIYIIGADNACMIPGNEEDIKNGPLRPKKCYLRFK
gi|12621078| NPPTCEGKTKFSGNEEIIYIIGADNACMIPGNEEDIKNGPLRPKKCYLRFK
gi|125977| DDYHNFTRNKKLEREMRNRFVRAVVDTPCK--LYTSSPFSEFLDLMRE
gi|10728878| DDYHNFTRNKKLEREMRNRFVRAVVDTPCK--LYTSSPFSEFLDLMRE
gi|7290546| PFLTSGSRKSSLEAEFTTIGTANCKHOCAG--YCNGPLRAGTIVYRIKI
gi|1362625| PFLTSGSRKSSLEAEFTTIGTANCKHOCAG--YCNGPLRAGTIVYRIKI

1910 1920 1930 1940 1950

NOV2A RATNIMCGFTDSSEYSDPEKTLGEGLSERIVELISVTLCLMSILDTGTAL
gi|12621078| RATNVMCGFTDSSEYSDPEKTLGEGLSERIVELISVTLCLMSILDTGTAL
gi|125977| APPGGERPHRPDPNAPPEPEVSNNRNKEPEIIVVWVPLAMSHFVSTALAI
gi|10728878| APPGGERPHRPDPNAPPEPEVSNNRNKEPEIIVVWVPLAMSHFVSTALAI
gi|7290546| RAFTLEDKFTDVIYSSPETTIE-----RSDTVIVRATVSAVLKAVVLA
gi|1362625| RAFTLEDKFTDVIYSSPETTIE-----RSDTVIVRATVSAVLKAVVLA

1960 1970 1980 1990 2000

5

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

FAPARIRKQ-----KEGGTYSQDAEIDDTKKLQDLITVADLELKRDER
FAFVRIRKQ-----KEGGTYSQDAEIDDTKKLQDLITVADLELKRDER
VLCVVRRRRRPOCKTPDOAAVTRPLMAADLGAGPTSPSPMDMRRINFTQPG
VLCVVRRRRRPOCKTPDOAAVTRPLMAADLGAGPTSPSPMDMRRINFTQPG
VYCOHFCOLI-----BRASKLAR-----MODELAAAPPEGYITPN-
VYCOHFCOLI-----BRASKLAR-----MODELAAAPPEGYITPN-

10

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2010 2020 2030 2040 2050

IT-----RPISKKSFLCHVBSLCINNNLKFCSEFSELPKFLQDLSS
ITRLLSYRKSISKISKKSFLCHVBSLCINNNLKFCSEFSELPKFLQDLSS
NIS-----HPPITPSEFANHIIRLKSNDNOKFSCYEYESTEPG-QQFTW
NIS-----HPPITPSEFANHIIRLKSNDNOKFSCYEYESTEPG-QQFTW

15

gi|7290546|
gi|1362625|

RPVAVKDSSEHYRIMSADSDFRFSEFEELKHVGRDOAC
RPVAVKDSSEHYRIMSADSDFRFSEFEELKHVGRDOAC

20

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2060 2070 2080 2090 2100

ITADLPWNRRAKNRFPNIKPYNNNRVKLIADASVPGSDYINASYILSCYLCP
ITADLPWNRRAKNRFPNIKPYNNNRVKLIADVSHPGSDYINASYILSCYLCP
DNSNLEHNKSKNRYANVTAYDHSRVQLPAVEGVGSDYINANYCDGYRKH
DNSNLEHNKSKNRYANVTAYDHSRVQLPAVEGVGSDYINANYCDGYRKH
SFANLPCNRPKNRFINILPYDHSRFLQPVDDDDGSDYINANYPCGHNPS
SFANLPCNRPKNRFINILPYDHSRFLQPVDDDDGSDYINANYPCGHNPS

25

gi|1362625|

30

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2110 2120 2130 2140 2150

NEFIATOGPLPGTVGDFWRMVMETRTKTLVMLTCFEKGRIRCHQYWPED
NEFIATOGPLPGTVGDFWRMVMETRTKTLVMLTCFEKGRIRCHQYWPED
NAYVATOGPLOETFVDFWRMCWELKTATIVMTRLEERTRIKCOQYWPTR
NAYVATOGPLOETFVDFWRMCWELKTATIVMTRLEERTRIKCOQYWPTR
REFIVTQGPLHSTREBFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPVC
REFIVTQGPLHSTREBFWRMCWESNSRAIVMLTRCFEKGREKCDQYWPVC

35

gi|1362625|

40

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2160 2170 2180 2190 2200

NKPVTVFGDVIITKLMEDVQIDWTITRDLKIERE--GDCMTVROCNETGWP
NKPVTVFGDVIITKLMEDVQIDWTITRDLKIERE--GDCMTVROCNETGWP
G--TETYGQIPVTTITETQELATYSIRTFQICROGFNDREIKCLOFTAMP
G--TETYGQIPVTTITETQELATYSIRTFQICROGFNDREIKCLOFTAMP
R-VAMFYGDIKVLQIITDTHYDWSISEFMYSRN--CESRIIMRHFFETTPW
R-VAMFYGDIKVLQIITDTHYDWSISEFMYSRN--CESRIIMRHFFETTPW

45

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2210 2220 2230 2240 2250

DHGVPENSAPIITHFVLVLRASRAHDTITPVHCSAGVGRGTGVFIALDHIT
DHGVPENITLHFVVLVLRASRAHDTITPVHCSAGVGRGTGVFIALDHIT
DHGVPDHPAEPFOFLRRCRAETPPESGPPVIVHCSAGVGRGTGVFIALDSML
DHGVPDHPAEPFOFLRRCRAETPPESGPPVIVHCSAGVGRGTGVFIALDSML
DFGVPEPPQSLVRFVRAFRDVIIGDMRPIIVHCSAGVGRSGTFIALDRIL
DFGVPEPPPLSLVRFVRAFRDVIIGDMRPIIVHCSAGVGRSGTFIALDRIL

50

gi|1362625|

55

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2260 2270 2280 2290 2300

QHINHDHFVDIYGLVABLRSERMCMVONLAQYIFIHQCILDILLS-----
QHINHDHFVDIYGLVABLRSERMCMVONLAQYIFIHQCILDILLS-----
ERVKHEKIIDYGHVTCRLACRNNMVOTEDQYIFIHDAILEAMICG---
ERVKHEKIIDYGHVTCRLACRNNMVOTEDQYIFIHDAILEAMICG---
QHIFKSDYVDIGCTVFAIRKERVVMVOTEDQYVCIHQCLLAFLLEGKEHLL
QHIFKSDYVDIGCTVFAIRKERVVMVOTEDQYVCIHQCLLAFLLEGKEHLL

60

gi|1362625|

65

NOV2A
gi|12621078|
gi|125977|
gi|10728878|
gi|7290546|
gi|1362625|

2310 2320 2330 2340 2350

-----NKCSNQPICEYNYSALQKVDSEDAAMEGGDVEEWESEITM
-----NKCGHQPVCEYNYSTLOKVDSEDAAMEG-DVEEWESEITM
--VTEUPARNLTHLQKLLITEPGETISGHEVEFKKLSNNKQSSKFVTA
--VTEUPARNLTHLQKLLITEPGETISGHEVEFKKLSNNKQSSKFVTA
ADSLERHANBCEVTHILYLRSPOTKNGPIPIRASLASAEKCDADLMTNK
ADSLERHANBCEVTHILYLRSPOTKNGPIPIRASLASAEKCDADLMTNK

		2360	2370	2380	2390	2400
5	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	NLPCNKHNNLVHLLPYESSRVYLTPIHGIEGSDYVNASFIDGYRYSAY				
	gi 10728878	NLPCNKHNNLVHLLPYESSRVYLTPIHGIEGSDYVNASFIDGYRYSAY				
	gi 7290546	DEDEDDEEQQQQQQ-----LATEVMPKGSN				
10	gi 1362625	DEDEDDEEQQQQQQ-----LATEVMPKGSN				
		2410	2420	2430	2440	2450
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
15	gi 125977	IAAQGPVQAAECFWRLWEHNSITIVMLTKLKEMGREKCFQYWPHERSV				
	gi 10728878	IAAQGPVQAAECFWRLWEHNSITIVMLTKLKEMGREKCFQYWPHERSV				
	gi 7290546	DDEEDDEEDDDDDQQLNNETTATUSSASCSSS-----THQVHV				
	gi 1362625	DDEEDDEEDDDDDQQLNNETTATUSSASCSSS-----THQVHV				
20		2460	2470	2480	2490	2500
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	RYQYYVVDIAEYNMPQYKLREFKVIDARDGSSRTVRQFQFIDWPEQVVP				
25	gi 10728878	RYQYYVVDIAEYNMPQYKLREFKVIDARDGSSRTVRQFQFIDWPEQVVP				
	gi 7290546	VLQEAIEKQEQERICAGTQSHADTESDNTDSDDDDEGDGKVAKDCAV				
	gi 1362625	VLQEAIEKQEQERICAGTQSHADTESDNTDSDDDDEGDGKVAKDCAV				
30		2510	2520	2530	2540	2550
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	KSGECFIDFIGVHKTKEQFGQDGPITVHCSAGVGRSGVPITLSIVLERM				
	gi 10728878	KSGECFIDFIGVHKTKEQFGQDGPITVHCSAGVGRSGVPITLSIVLERM				
35	gi 7290546	ADEDCWY-----				
	gi 1362625	ADEDCWY-----				
40		2560	2570	2580	2590	
	NOV2A				
	gi 12621078	----- ----- ----- ----- ----- ----- -----				
	gi 125977	QYEGVLDVFQTVRILRSQRPAMVQTEDQYHFCYRAALEYLGSFDNYTN				
	gi 10728878	QYEGVLDVFQTVRILRSQRPAMVQTEDQYHFCYRAALEYLGSFDNYTN				
	gi 7290546	----- ----- ----- ----- ----- ----- -----				
45	gi 1362625	----- ----- ----- ----- ----- ----- -----				

Tables 2J-2EE list the domain descriptions from DOMAIN analysis results against NOV2a. This indicates that the NOV2a sequence has properties similar to those of other proteins known to contain this domain.

50

Table 2J. Domain Analysis of NOV2a

gnl|Smart|smart00194, PTPc, Protein tyrosine phosphatase, catalytic domain (SEQ ID NO:93)
 CD-Length = 264 residues, 99.6% aligned
 Score = 318 bits (816), Expect = 2e-87

55

NOV 1: 1983 KFQEEFSELPK-FLQDLSSDADLPWNRAKNRFPNIKPYNNRVKLIADASVPGSDYINA 2041
 +||| +| + ||| | | | | | | + + || + ||| |||||
 Sbjct: 1 GLEEEFEKLQRLTPDDLSCVTAILPENRDKNRYKDVLPYDHTRVKL-KPPPGEGSDYINA 59
 NOV 1: 2042 SYISGYLCPNEFIATQGPLPGTVGDFWRMVWETRAKTLVMLTQCFEKGRIRCHQYWPEDN 2101

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 3D.

Table 3D. ClustalW Analysis of NOV3

5
1) NOV3 (SEQ ID NO:12)
2) ref|XP_027243.1| (XM_027243) hypothetical protein XP_027243 [Homo sapiens] (SEQ ID NO:42)
3) gi|15076843|gb|AAK82958.1|AF233450_1 (AF233450) pecanex-like protein 1 [Homo sapiens] (SEQ ID NO:43)
10 4) gi|6650377|gb|AAF21809.1|AF096286_1 (AF096286) pecanex 1 [Mus musculus] (SEQ ID NO:44)
5) gi|13171105|gb|AAK13590.1|AF154413_1 (AF154413) pecanex [Takifugu rubripes] (SEQ ID NO:45)
15 6) gi|7290294|gb|AAF45755.1| (AE003423) pcx gene product [alt 1] [Drosophila melanogaster] (SEQ ID NO:46)

20

NOV3
ref|XP_027243.1
gi|15076843|gb|MGSQTLQILRQGVWAALSGGWYDPHQATFVNALHLYLWLFLGLPFTLYMALPSTMIIV 60
gi|6650377|gb|A-----
gi|13171105|gb|MGSQTLQILRQGVWASVTGGWYDDPDQNTFVNALHLYIWLFLLCFPFTLYMALQPSMVIV 60
gi|7290294|gb|A-----

25

70 80 90 100 110 120
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
NOV3-----1
ref|XP_027243.1-----1
gi|15076843|gb|AVYCPVIAAVFIVLKMVNYRLHRALDAGEVVDRTANEFTDQR-TKAEQGNCSTRRKDSNG 119
gi|6650377|gb|A-----1
gi|13171105|gb|GIYCGVIAAMFLLLKTVNYRLHHALDEGEVVEHQTRESKSRGGTGGANDPVTRREDSNG 120
gi|7290294|gb|A-----1

30

130 140 150 160 170 180
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
NOV3-----1
ref|XP_027243.1-----1
gi|15076843|gb|PSDPGGGIEMSEFIREATPPVGCSSRSNSYAGLDPSNQIGSGSRLGTAATIKGDTDTAKT 179
gi|6650377|gb|A-----1
gi|13171105|gb|LGDPGGGIEMADFIROETPPVDCSSRSNSYVG-----151
gi|7290294|gb|A-----1

35

190 200 210 220 230 240
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
NOV3-----1
ref|XP_027243.1-----1
gi|15076843|gb|SDDISLSLQGSSSLCKEGSEEQLDAADRKLPLRVSNDSFISIQPSLSSCGQDLPRDFSDK 239
gi|6650377|gb|A-----1
gi|13171105|gb|-----151
gi|7290294|gb|A-----1

40

250 260 270 280 290 300
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
NOV3-----1
ref|XP_027243.1-----1
gi|15076843|gb|VNLPSHNHHHVVDQSLSSACDTEVASLVPLHSHSYRKDHRPRGVPTSSSAVAFPDTSLN 299
gi|6650377|gb|A-----1
gi|13171105|gb|-----151
gi|7290294|gb|A-----1

45

310 320 330 340 350 360
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
NOV3-----1
ref|XP_027243.1-----1

50

56



5	NOV3	-----	1
	ref XP_027243.1	-----	1
	gi 15076843 gb	SFRRERSTFRRQAVRRRHNAGSNPTPTLLIGSPLSLQDGGQQQSTAQ-----VKVQS	833
	gi 6650377 gb A	-----	1
	gi 13171105 gb	IFRRERSTFRRQAVRRRHNAGSNPTPTSLIGSPLRYALHEADRPSGVRSWYRTVKSQPS	237
	gi 7290294 gb A	-----	1
10	NOV3	850 860 870 880 890 900	3
	ref XP_027243.1	-----	1
	gi 15076843 gb	RPPSQAAVLSASASLLVRNGSVHLEASHDNASAVGGSSSLHDELGKFSSTLYETGCGDMSL	893
	gi 6650377 gb A	-----	3
	gi 13171105 gb	RTPSQVTVLSTSASLLARNGSTHLEGSQDKASTVGTTSLQDEFGLTTPSLYEIRGCHGCL	297
	gi 7290294 gb A	-----	1
20	NOV3	910 920 930 940 950 960	57
	ref XP_027243.1	-----	1
	gi 15076843 gb	VNFEPARRASN-ICDTSHSVSSSTSVRFYPHDVTL-----RLNRLLTIDTLLLEQQDIDL	952
	gi 6650377 gb A	VNFEPARRASN-ICDTSHSVSSSTSVRFYPHDVTL-----LNRLLTIDTLLLEQQDIDL	57
	gi 13171105 gb	CNFESARRASN-ICDTSHSVSSSTSVRFYPHDVTL-----LNRLLTIDTLLLEQQDIDL	356
	gi 7290294 gb A	-----	1
30	NOV3	970 980 990 1000 1010 1020	112
	ref XP_027243.1	-----	1
	gi 15076843 gb	SPDLAAT-----YGPTEEAACRVKHYYRFWILPOLWICINFDRLTLLALFDRNREILENV	1007
	gi 6650377 gb A	SPDLAAT-----YGPTEEAACRVKHYYRFWILPOLWICINFDRLTLLALFDRNREILENV	112
	gi 13171105 gb	SPDLQDAPLGQDNPSAASAAGTROYVRLWILPFLWVGLHFDRLTLLALFDRNREILENV	416
	gi 7290294 gb A	-----	1
40	NOV3	1030 1040 1050 1060 1070 1080	172
	ref XP_027243.1	-----	1
	gi 15076843 gb	LAVILAILVAFVLSILLIGFFRDIWVQFCLVIASCQYSLKSVQPDSSSPRHGHNRIL	1067
	gi 6650377 gb A	LAVILAILVAFVLSILLIGFFRDIWVQFCLVIASCQYSLKSVQPDSSSPRHGHNRIL	172
	gi 13171105 gb	LAVVLAIVLAFVLSVLLIGFFRDIWVQFCLVIASCQYSLKSVQPDSSSPRHGHNRIL	476
	gi 7290294 gb A	-----	1
50	NOV3	1090 1100 1110 1120 1130 1140	232
	ref XP_027243.1	-----	1
	gi 15076843 gb	AYSRPVYFCICCGLIWLLDYGSRNLITATKFKLYGTFITNPLVFISARDLVIVFTLCFPIV	1127
	gi 6650377 gb A	AYSRPVYFCLCCGLIWLLDYGSRNLITATKFKLYGTFITNPLVLLSARDLVIVFTLCFPIV	232
	gi 13171105 gb	AYSRPVYFCLCCGLIWLLHYGSLRTISRFITLYGVALTSSLVLASARDLVIVFTLCFPIV	536
	gi 7290294 gb A	-----	1
60	NOV3	1150 1160 1170 1180 1190 1200	292
	ref XP_027243.1	-----	1
	gi 15076843 gb	FFIGLLPQVNTFVMYLCQOLDIHIFGGNATISLLAALYSFICSIVAVALLYGLCYGALKQ	1187
	gi 6650377 gb A	FFIGLLPQVNTFVMYLCQOLDIHIFGGNATISLLAALYSFICSIVAVALLYGLCYGALKQ	292
	gi 13171105 gb	FFVGLLPQVNTFVMYLCQOLDIHIFGGNASTSLLAALYSTIRSIVTALLYCFCYGALKK	596
	gi 7290294 gb A	-----	1
70	NOV3	1210 1220 1230 1240 1250 1260	352
	ref XP_027243.1	-----	1
	gi 15076843 gb	SMGCOHHPVLFSIFCGLLVAVSYHLSROSSDPSVLSIVOSKIFPKTEERKNPEDPLSEVK	1247
	gi 6650377 gb A	SMGCOHHPVLFSIFCGLLVAVSYHLSROSSDPSVLSIVOSKIFPKTEERKNPEDPLSEVK	352
	gi 13171105 gb	SMGPHHHPVLFSIFCGLLVAVSYHLSROSSDPSVLSIVOSKIFPKTEERKNPEDPLSEVK	656
	gi 7290294 gb A	-----	1

		1270	1280	1290	1300	1310	1320	
5	NOV3	DPLPEKLRNSVSRERLQSDLVVCTVIGVLYFAIHVSTVFTVLOPALKYVLYTLVGVGVFT	412					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	DPLPEKLRNSVSRERLQSDLVVCTVIGVLYFAIHVSTVFTVLOPALKYVLYTLVGVGVFT	1307					
	gi 6650377 gb A	DPLPEKLRNSVSRERLQSDLVVCTVIGVLYFAIHVSTVFTVLOPALKYVLYTLVGVGVFT	412					
	gi 13171105 gb	DPLPEKLRASVNERLQSDLVVCTVIGVLYFAIHVSTVFTVLOPALKYVLYTLVGVGVFT	716					
	gi 7290294 gb A	-----	1					
10		1330	1340	1350	1360	1370	1380	
	NOV3	HYVLPQVRKQLPWHCFSPHLLKTLBYNQEVRDAATMMWFELHVVLLFVEKNIIYPLIV	472					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	HYVLPQVRKQLPWHCFSPHLLKTLBYNQEVRDAATMMWFELHVVLLFVEKNIIYPLIV	1367					
15	gi 6650377 gb A	HYVLPQVRKQLPWHCFSPHLLKTLBYNQEVRDAATMMWFELHVVLLFVEKNIIYPLIV	472					
	gi 13171105 gb	HYVLPQVRKQLPWHCFSPHLLKTLBYNQEVRDAATMMWFELHVVLLFVEKNIIYPLIV	776					
	gi 7290294 gb A	-----	1					
20		1390	1400	1410	1420	1430	1440	
	NOV3	LNELSSAETIASPKRLNTELGALMITVAGLKLLRSSSSSPTYQYITVIFTVLFFKFDYE	532					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	LNELSSAETIASPKRLNTELGALMITVAGLKLLRSSSSSPTYQYITVIFTVLFFKFDYE	1427					
	gi 6650377 gb A	LNELSSAETIASPKRLNTELGALMITVAGLKLLRSSSSSPTYQYITVIFTVLFFKFDYE	532					
25	gi 13171105 gb	LNELSSAETIASPKRLNTELGALMITVAGLKLLRSSSSSPTYQYITVIFTVLFFKFDYE	836					
	gi 7290294 gb A	-----	1					
30		1450	1460	1470	1480	1490	1500	
	NOV3	AFSETMLLDLEFMSITLFFKRWELLYKLOFVYTYIAPWQITWGSFAHFAQPPFAVPHSAML	592					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	AFSETMLLDLEFMSITLFFKRWELLYKLOFVYTYIAPWQITWGSFAHFAQPPFAVPHSAML	1487					
	gi 6650377 gb A	AFSETMLLDLEFMSITLFFKRWELLYKLOFVYTYIAPWQITWGSFAHFAQPPFAVPHSAML	592					
	gi 13171105 gb	HLSETMLLDLEFMSITLFFKRWELLYKLOFVYTYIAPWQITWGSFAHFAQPPFAVPHSAML	896					
35	gi 7290294 gb A	-----	1					
40		1510	1520	1530	1540	1550	1560	
	NOV3	FVQAIVSAFFSTPLNPFGLSAIFITSYVRPVKFWERDYNTKRVDHSNTRLASQDRNPQS	652					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	FVQAIVSAFFSTPLNPFGLSAIFITSYVRPVKFWERDYNTKRVDHSNTRLASQDRNPQS	1547					
	gi 6650377 gb A	FVQAIVSAFFSTPLNPFGLSAIFITSYVRPVKFWERDYNTKRVDHSNTRLASQDRNPQS	652					
	gi 13171105 gb	FVQAIVSAFFSTPLNPFGLSAIFITSYVRPVKFWERDYNTKRVDHSNTRLASQDRNPQS	956					
45	gi 7290294 gb A	-----	47					
		MSTEPSSEDEXT-SAPPVDCRYTDLK-----ENEMKQVDFEDTRVLLMKON----						
50		1570	1580	1590	1600	1610	1620	
	NOV3	DDNNLNSIFYEHLTRSLQHSLSLGGDLLGRWGNVSTGDCFILASDYLNALVHLIEIGNGLV	712					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	DDNNLNSIFYEHLTRSLQHSLSLGGDLLGRWGNVSTGDCFILASDYLNALVHLIEIGNGLV	1607					
	gi 6650377 gb A	DDNNLNSIFYEHLTRSLQHSLSLGGDLLGRWGNVSTGDCFILASDYLNALVHLIEIGNGLV	712					
	gi 13171105 gb	DDNNLNSIFYEHLTRSLQHSLSLGGDLLGRWGNVSTGDCFILASDYLNALVHLIEIGNGLV	1016					
	gi 7290294 gb A	DRLLAVGAKCTHYGAPLOT---CALGLGRVRCPHWGACPNLENGDIEDFP-----GLD	97					
55		1630	1640	1650	1660	1670	1680	
	NOV3	TFQLRGLFRGTTCQOREVEAITEGVEEDEGFCCEPGHHPHLSFNAAFQORWLAWEV	772					
	ref XP_027243.1	-----	1					
	gi 15076843 gb	TFQLRGLFRGTTCQOREVEAITEGVEEDEGFCCEPGHHPHLSFNAAFQORWLAWEV	1667					
60	gi 6650377 gb A	TFQLRGLFRGTTCQOREVEAITEGVEEDEGFCCEPGHHPHLSFNAAFQORWLAWEV	772					
	gi 13171105 gb	TFQLRGLFRGTTCQOREVEAITEGVEEDEGFCCEPGHHPHLSFNAAFQORWLAWEV	1076					
	gi 7290294 gb A	SLP-----	105					
65		1690	1700	1710	1720	1730	1740	
	NOV3	VTKYILEGYSITDNSAASMLQVFLRRVLTYYVKGIIYVVTSSKLEEWLANETMQEGL	832					
	ref XP_027243.1	-----	5					
	gi 15076843 gb	VTKYILEGYSITDNSAASMLQVFLRRVLTYYVKGIIYVVTSSKLEEWLANETMQEGL	1727					
	gi 6650377 gb A	VTKYILEGYSITDNSAASMLQVFLRRVLTYYVKGIIYVVTSSKLEEWLANETMQEGL	832					
70	gi 13171105 gb	VTKYILEGYSITDNSAASMLQVFLRRVLTYYVKGIIYVVTSSKLEEWLANETMQEGL	1136					

	gi 6650377 gb A	GFVPCRRSSTSQISLRNLPSSIQSRLSMVNQMEARSGGVCVOHCLPSSSSSSQSIAC	1308
	gi 13171105 gb	GLEPCRRSSTSQISLRNLPSTQLRLGST--SDPAG-----PSSSTSSHSTPFC	1558
	gi 7290294 gb A	--G-----VKKLEAFPFFFTLLFGKG--RRYAC-----HG--SYKDVIIDGEM	485
5		2230 2240 2250 2260 2270 2280	
	NOV3	KHHTLVGFLATEGGQSSATDAC-----PGNTLSPANNSHS--RKAEEVIYRVOIVDFSOIL	1361
	ref XP_027243.1	KHHTLVGFLATEGGQSSATDAC-----PGNTLSPANNSHS--RKAEEVIYRVOIVDFSOIL	534
	gi 15076843 gb	KHHTLVGFLATEGGQSSATDAC-----PGNTLSPANNSHS--RKAEEVIYRVOIVDFSOIL	2256
10	gi 6650377 gb A	KHHTLVAFLEGGQSSATEAC-----PGNTLSPANNISHA--RKGEVIYRVOIVDFSOIL	1361
	gi 13171105 gb	KHHTLVGFLGNDGLCSVTCTPLSQHHHPHHHPQHNPHTATVRDDISYRVOIVDFGOVL	1618
	gi 7290294 gb A	EDFKFVAFFINEADTVLAVASC-----G-----RDPVIAQFAELTSQGXCL	526
15		2290 2300 2310 2320 2330 2340	
	NOV3	EGINLSKRKELQWPDEGIRLKAGRNSWKDWSPQEGMEGHVTHRWVPCSRDPGTRSHIDKA	1421
	ref XP_027243.1	EGINLSKRKELQWPDEGIRLKAGRNSWKDWSPQEGMEGHVTHRWVPCSRDPGTRSHIDKA	594
	gi 15076843 gb	EGINLSKRKELQWPDEGIRLKAGRNSWKDWSPQEGMEGHVTHRWVPCSRDPGTRSHIDKA	2316
	gi 6650377 gb A	EGINLSKRKELQWPDEGIRLKAGRNSWKDWSPQEGMEGHVTHRWVPCSRDPGTRSHIDNT	1421
20	gi 13171105 gb	EGINLSKRKELQWPDEGIRLKAGRNSWKDWSPQEGMEGHVTHRWVPCSRDPGTRSHIDKT	1678
	gi 7290294 gb A	G-----RQIEEDP-----A-----RRPDWTKKLCQP-----LPQVR-----	552
25		2350 2360	
	NOV3	VLLVQIDDKYVTIETGVLELGAEV	1446
	ref XP_027243.1	VLLVQIDDKYVTIETGVLELGAEV	619
	gi 15076843 gb	VLLVQIDDKYVTIETGVLELGAEV	2341
	gi 6650377 gb A	VLLVQIDDKYVTIETGVLELGAEV	1446
	gi 13171105 gb	VLLVQVDDKTYVTIETGVLELGAEV	1703
30	gi 7290294 gb A	-----	552

Pecanex gene was originally discovered in *Drosophila*, encoding a large, membrane-spanning protein. The mouse homolog was recently reported. In the absence of maternal expression of the pecanex gene, the embryo develops severe hyperneuralization similar to that characteristic of Notch mutant embryos. Early gastrula embryos, lacking both maternally and zygotically expressed activity of the neurogenic pecanex locus, are shown to contain a greater than wild-type number of stably determined neural precursor cells which can differentiate into neurons in culture. Therefore it is anticipated that this novel human pecanex will be involved in neuronal differentiation, maintenance of neuronal precursors and neurological diseases.

The disclosed NOV3 nucleic acid of the invention encoding a Human homolog of the *Drosophila* pecanex protein includes the nucleic acid whose sequence is provided in Table 3A or a fragment thereof. The invention also includes a mutant or variant nucleic acid any of whose bases may be changed from the corresponding base shown in Table 3A while still encoding a protein that maintains its Human homolog of the *Drosophila* pecanex activities and physiological functions, or a fragment of such a nucleic acid. The invention further includes nucleic acids whose sequences are complementary to those just described, including nucleic acid fragments that are complementary to any of the nucleic acids just described. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications. Such modifications include, by way

gi 13653970 ref XP_009546.3 (XM_009546)	serine/threonine kinase 15 [Homo sapiens]	403	369/403 (91%)	381/403 (93%)	0.0
gi 4507275 ref NP_03591.1 (NM_003600)	serine/threonine kinase 15; Serine/threonine protein kinase 15 [Homo sapiens]	403	369/403 (91%)	380/403 (93%)	0.0
gi 7446411 pir JC5974	aurora-related kinase 1 (EC 2.7.--.) - human	403	367/403 (91%)	379/403 (93%)	0.0
gi 4507279 ref NP_03149.1 (NM_003158)	serine/threonine kinase 6; Serine/threonine protein kinase-6; serine/threonine kinase 6 (aurora/IPL1-like) [Homo sapiens]	402	342/403 (84%)	360/403 (88%)	0.0

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 4D.

5

Table 4D. ClustalW Analysis of NOV4

- 1) NOV4 (SEQ ID NO:14)
- 2) gi|12654873|gb|AAH01280.1|AAH01280 (BC001280) serine/threonine kinase 15 [Homo sapiens] (SEQ ID NO:47)
- 3) gi|13653970|ref|XP_009546.3| (XM_009546) serine/threonine kinase 15 [Homo sapiens] (SEQ ID NO:48)
- 4) gi|4507275|ref|NP_03591.1| (NM_003600) serine/threonine kinase 15; Serine/threonine protein kinase 15 [Homo sapiens] (SEQ ID NO:49)
- 5) gi|7446411|pir|JC5974 aurora-related kinase 1 (EC 2.7.--.) - human (SEQ ID NO:50)
- 6) gi|4507279|ref|NP_03149.1| (NM_003158) serine/threonine kinase 6; Serine/threonine protein kinase-6; serine/threonine kinase 6 (aurora/IPL1-like) [Homo sapiens] (SEQ ID NO:51)

20

25

30

35

		10	20	30	40	50
NOV4	MD	SE	ENCIS	GPV	SA	KTPVGGP
gi 12654873	MD	SK	ENCIS	GPV	KATAPVGGP	KRV
gi 13653970	MD	SK	ENCIS	GPV	KATAPVGGP	KRV
gi 4507275	MD	SK	ENCIS	GPV	KATAPVGGP	KRV
gi 7446411	MD	SK	ENCIS	GPV	KATAPVGGP	KRV
gi 4507279	MD	SK	ENCIS	GPV	KATAPVGGP	KRV
		60	70	80	90	100
NOV4	SN	SS	Q	R	V	P
gi 12654873	SN	SS	Q	R	V	P
gi 13653970	SN	SS	Q	R	V	P
gi 4507275	SN	SS	Q	R	V	P
gi 7446411	SN	SS	Q	R	V	P

[illegible]

NOV5	RKIEFPLPDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI
gi 4506207	RKIEFPLPDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI
gi 6679501	RKIEFPLPDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI
gi 345717	RKIEFPLPDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI
5 gi 16741033	RKIEFPLPDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI
gi 2492516	RKIEFPLPDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI

	410	420	430	440
NOV5	CTEAGLMALRERMKVTNEDFKKSTESVLYKKKEGTPEGLYL			
gi 4506207	CTEAGLMALRERMKVTNEDFKKSKENVLYKKQEGTPEGLYL			
gi 6679501	CTEAGLMALRERMKVTNEDFKKSKENVLYKKQEGTPEGLYL			
gi 345717	CTEAGLMALRERMKVTNEDFKKSKENVLYKKQEGTPEGLYL			
gi 16741033	CTEAGLMALRERMKVTNEDFKKSKENVLYKKQEGTPEGLYL			
15 gi 2492516	CTEAGLMALRERMKVTNEDFKKSKENVLYKKQEGTPEGLYL			

Tables 5E-F list the domain description from DOMAIN analysis results against NOV5.

This indicates that the NOV5 sequence has properties similar to those of other proteins known to contain this domain.

Table 5E. Domain Analysis of NOV5

gnl|Pfam|pfam00004, AAA, ATPase family associated with various cellular activities (AAA). AAA family proteins often perform chaperone-like functions that assist in the assembly, operation, or disassembly of protein complexes (SEQ ID NO:101)
CD-Length = 186 residues, 100.0% aligned
Score = 190 bits (483), Expect = 1e-49

25	NOV 4:	221	GVILCGPPGTGKTLAKAVANQTSATFLRVVGS	ELIQYLG	DGPKLVRQV	FQVAEEHAPS	280		
	Sbjct:	1	GILLYGPPGTGKTLAKAVAKELGVPFIEISG	SELLSKYVGESEKLV	RALFSLARKSAPC		60		
30	NOV 4:	281	IMFTDEIEAIGTKRYDSNSGGEREIQOTM	LELELLN	QLCGFDS	REDVKVIMATKQVETLD	340		
	Sbjct:	61	IIFIDEIDALAPKRGDVGTDVSS---	RVVNQLL	TEM	DGFEKLSNVIVIGATNRPDLLD	116		
35	NOV 4:	341	PVLIRPGRIDKKIEFHL	PDEKTKKRIFQIH	TSRMTL	ADVTLD	DLIMAKDD	SGADIKAI	400
	Sbjct:	117	PALLRPGRFDRRIE	VPLPDEEERLEIL	KIHLKKKPLEK	DVLD	DEIARRTPGFS	GDADLAAL	176
35	NOV 4:	401	CTEAGLMALR	410					
	Sbjct:	177	CREAALRAIR	186					

Table 5F. Domain Analysis of NOV5

gnl|Smart|smart00382, AAA, ATPases associated with a variety of cellular activities; AAA. This profile/alignment only detects a fraction of this vast family. The poorly conserved N-terminal helix is missing from the alignment. (SEQ ID NO:102)
CD-Length = 151 residues, 100.0% aligned
Score = 61.6 bits (148), Expect = 9e-11

40	NOV 4:	218	PPKGVILCGPPGTGKTLAKAVANQTSATFLRVV-----GSELIQK	258
			+ ++ + + + + +	
	Sbjct:	1	PGEVVLIVGPPGSGKTTLRALLARELGPDGGGVYIDGEDLREEALLQLRLLLVLVGEDK	60

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 6D.

Table 6D Clustal W Sequence Alignment

5	1) NOV6 (SEQ ID NO:18)
	2) gi 3077703 dbj BAA25784.1 (AB004816) mitsugumin29 (<i>Oryctolagus cuniculus</i>) (SEQ ID NO:57)
	3) gi 6678874 ref NP_032622.1 (NM_008596) mitsugumin 29 (<i>Mus musculus</i>) (SEQ ID NO:58)
10	4) gi 12836843 dbj BAB23831.1 (AK005132) putative (<i>Mus musculus</i>) (SEQ ID NO:59)
	5) gi 1351168 sp P20488 SYPH_BOVIN SYNAPTOPHYSIN (MAJOR SYNAPTIC VESICLE PROTEIN P38) (SEQ ID NO:60)
	6) gi 2134413 pir I50720 synaptophysin IIa - chicken (SEQ ID NO:61)
15	
	<div> <div> 1020304050 </div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> MSSTESAGRTADKSPRQVDRLLVGLRWRLLEPLGFLKVLQWLFAIFAF MSSTESPSRAADKSPRQVDRLLVGLRWRLLEPLGFLKVLQWLFAIFAF MSSTESPGRTSDKSPRQVDRLLVGLRWRLLEPLGFLKVLQWLFAIFAF -----MDPVSVQASAGTFRALKEPLAFLRALFLFAIFAF -----MDVNVQLVAGGQERVWKEPLGFLKVLQWLFAIFAF -----MCVVLFAPLFAIFAF </div> </div>
20	
	<div> <div>60708090100</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> GSCGSYSGETGAMVRONNEAKDVSSIIIVAFGYPPFRLHRTQVEMFLQDEES GSCGSYSGETGAMVRONNEAKDVSSIIIVAFGYPPFRLHRTQVEMFLQDDDS GSCGSYSGETGALVLONNEAKDVSSIIIVAFGYPPFRLHRTQVEMFLQDDDS ATCGGYSGLRLSVDCANKKTSNLSDIAFAVPPFRLQVTFEVPFQSGK- ATCGSYSGLQLSVDCANKKTSNLSDIAFAVPPFRLHRTQVTFEVPFQSGK- ATCGGYSGLRLSVDCANKKTSNLSDIAFAVPPFRLHRTQVTFEVPFQSGK- </div> </div>
25	
	<div> <div>110120130140150</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> SSKTMHLMGDFSAPAEFFVTLGTFSSFYMAALVIVRFHKLTYTENKRFEP SSKTMHLMGDFSAPAEFFVTLGTFSSFYMAALVIVRFHKLTYTENKRFEP TSKTMHLMGDFSAPAEFFVTLGTFSSFYMAALVIVRFHKLTYTENKRFEP EQCKIALVGDSSSAEAEFFVTVAFPAFLYSIAATVVYVFFONKYRENNRGP DPKRIFLVGNYSSEAEFFVTVAFPAFLYSIAATVVYVFFONKYRENNRGP RRRTUSLIGDFSSEAEFFVTVAFPAFLYSIAATVVYVFFONKYRENNRGP </div> </div>
30	
	<div> <div>160170180190200</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> LVDFQVTVSFTFFWLVAANAAGKGLTDVKCATRPSSTTAAMSVCHGEAV LVDFQVTVSFTFFWLVAANAAGKGLTDVKCATRPSSTTAAMSVCHGEAV LVDFQVTVSFTFFWLVAANAAGKGLTDVKCATRPSSTTAAMSVCHGEAV LIDFIVTVVESFLWLVCSSAAKGLSDVKVATDPEKLLMSACKQPSNK MLDFLATAVFAFMWLVSAAKGLSDVKVATDPEKLLMSACKQPSNK LIDFIVTVVESFLWLVCSSAAKGLSDVKVATDPEKLLMSACKQPSNK </div> </div>
35	
	<div> <div>210220230240250</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 gi 2134413 </div> <div> CSAGATPSMGLANISVIFGFINFFLWAGNCFVFKETPNHGGCGGQDQDQ CSAGATPSMGLANISVIFGFINFFLWAGNCFVFKETPNHGGCGGQDQDQ CSAGATPSMGLANISVIFGFINFFLWAGNCFVFKETPNHGGCGGQDQDQ QWVHSPVMSSLNISVIFGFINFFLWAGNCFVFKETPNHGGCGGQDQDQ CKELRDPVTSGLNISVIFGFINFFLWAGNCFVFKETPNHGGCGGQDQDQ CLPVRSPVMSSLNISVIFGFINFFLWAGNCFVFKETPNHGGCGGQDQDQ </div> </div>
40	
	<div> <div>260270280290300</div> <div> NOV6 gi 3077703 gi 6678874 gi 12836843 gi 1351168 </div> <div> DQDQ-----GQGF-----SQESAEEQG --DQ-----GQGF-----SQESAEEQG --DQ-----GQGF-----SQESAEEQG MEKH-----SSSYNCG--RYN-QESYGSYGYS--QQAN--L PEKCPAPGDAYGQAGYGQGFGGYGPQESYGPQGYQPDYGPASSGGGY </div> </div>
45	
50	
55	
60	
65	

Table 6E. Domain Analysis of NOV6

gnl|Pfam|pfam01284, Synaptophysin, Synaptophysin / synaptoporin. (SEQ
ID NO:103)
CD-Length = 298 residues, 70.8% aligned
Score = 244 bits (622), Expect = 6e-66

	NOV 5:	29	RRLEELPLGFIKVLQWLFAIFAFGSCGSYSGETGAMVRCNNEAKDVSSIIIVAFGYPFRLHR	88
			+ + + + + + + + +	
20	Sbjct:	3	MVIFAPLPGFVKVLQWVFAIFAFATCGGYSGELQLSVDCAKNTESDLNIDIAFAYPFRLHE	62
	NOV 5:	89	IQYEMPLCDEESSKTMHLMGDFSAPAEFFVTLGIFSFYTYMAALVIYLRFHNLTYTENKR	148
			+ + + + + + + + + + + + +	
25	Sbjct:	63	VTFEAPTC-EGDEKQNIALVGDSSSSAEFFVTVAVFALYSLAALATYIFFQNKYRENNK	121
	NOV 5:	149	FPLVDFCVTVSFTFFWLVAANAAGKGLTDVKGATRPSSLTAAMSVCHGEEAVCSAGATPS	208
			+ ++ + +	
	Sbjct:	122	GPLIDFIATAVFAFLWLVGSSAWAKGLSDVKMATDPEEIIKGMHACHQPKNCKELHDPV	181
30	NOV 5:	209	MGLANISVLFGFINFFLWAGNCWFVPKETPWH	240
			+	
	Sbjct:	182	MSGLNTSVVFGFLNFIWLWAGNIWFPKETGWA	213

79

gi|16303264| ----MRPPPALALAGLQLLALPAARASVFGLTCRE----VLTIPFGL
gi|3915306| ----MALLRALLG--LLACTPRPSAIVFGLTCNE----ALTILP-L
gi|15082261| MLDGSPLARWLAAAFGLTLTLLAALRPSAIVFGLTCSE----PLTILP-L
gi|139748| ----MRILTFLGLKTLWVLAFFSLNTIAVNNSGKWWGIVNVASAG
gi|3024851| ----

5

60 70 80 90 100

NOV7
gi|16303264| GTAAAPAGGCAHLKQCDLLKLSRROKQLCRRPGLAETLRDAHGLLEC
gi|3915306| GTAAAPAGGCAHLKQCDLLKLSRROKQLCRRPGLAETLRDAHGLLEC
gi|15082261| TSEMEDAAVKAHYKVCRLKLEKKORRMCRDPGCAETLMETISYSALEC
gi|139748| TLEP-EAAACAHYKACDRLKLERKORRMCRDPGCAETLVEAVSYSALEC
gi|3024851| NVLPGSDARPVPLVDPSLQLLSRQKRIIRNPGILQSTIRGLHSATREC

10

110 120 130 140 150

NOV7
gi|16303264| QFQFRHERWNCSELEG--RMG--LLKRGFKETAFLYAVSSAALHTHTARAC
gi|3915306| QFQFRHERWNCSELEG--RTG--LLKRGFKETAFLYAVSSAALHTHTARAC
gi|15082261| QFQFRFERWNCLEGRYRAS--LLKRGFKETAFLYATSSAGLTHAMAKAC
gi|139748| QFQFRFERWNCLEGRYRAS--LLKRGFKETAFLYATSSAGLTHAMAKAC
gi|3024851| KWHFRNRWNCPTGTGNGVFGKIDNRGCRETAFVFAITSAGVTHSVARSC

15

160 170 180 190 200

NOV7
gi|16303264| SAGRMECTCDDSPGLESRQAWQWGVCGDNLKYSTKFTSNFLGSKRGNKD
gi|3915306| SAGRMECTCDDSPGLESRQAWQWGVCGDNLKYSTKFTSNFLGSKRGNKD
gi|15082261| SAGRMECTCDEADLENREAWQWGGCGDNLKYSNKFVKEFLG-RKPNKD
gi|139748| SAGRMECTCDEADLENREAWQWGGCGDNLKYSSKFVKEFLG-RRSSKD
gi|3024851| SEGSTESCSQDYRRRGPGGPDHFWGCSQNTIEGRFIGREFVDSSERGRD

20

210 220 230 240 250

NOV7
gi|16303264| LRARADAHNTHVGTNAVSGGRTTCKCHGVSGSCAVRTCWKQLSPFRETE
gi|3915306| LRARADAHNTHVGTNAVSGGRTTCKCHGVSGSCAVRTCWKQLSPFRETE
gi|15082261| LRARVDFHNNLVGKVIKAGVETTCCKCHGVSGSCTVRTCWQLSPFHEIC
gi|139748| LRARVDFHNNLVGKVIKAGVETTCCKCHGVSGSCTVRTCWQLSPFHEVE
gi|3024851| LKYLVLNHNNOAGRLTMTLTEMROECKCHMSGSCSLRTCWMRLPPFRSVG

25

260 270 280 290 300

NOV7
gi|16303264| QVLKLYRDSAVKVSSATNEALCRLELWAPAR---QGLTKGLAPRSCDLV
gi|3915306| QVLKLYRDSAVKVSSATNEALCRLELWAPAR---QGLTKGLAPRSCDLV
gi|15082261| KQLKQYETSLKVCSTTNEATGE-GDISPEK--KSIPGHSDQIPRTDLV
gi|139748| KHLKHKYETALKVCSTTNEAAGEAGATSPPRGRASCAGGSDPLPRPDELV
gi|3024851| DALKDRFDCASKVTHYSNNGSNRWGSRSDPPH--LAPENPTHALPSSODLV

30

310 320 330 340 350

NOV7
gi|16303264| YMEDSPSFCRPSKYS--PGTAGRVCSRE---ASCSSLCCGRGYDTQSRD
gi|3915306| YMEDSPSFCRPSKYS--PGTAGRVCSRE---ASCSSLCCGRGYDTQSRD
gi|15082261| YLDDSPSFCRLMSRYS--PGTSGRKQYKI---KNQDSICCGRGHNTQSRV
gi|139748| YLEKSPNFCSPSKNGTPGTIGRIQNSTSLGLDGCCLCCGRGYRSLAEK
gi|3024851| YMEDSPSFCRPSKYS--PGTAGRVCSRE---ASCSSLCCGRGYDTQSRD

35

360 370 380

NOV7
gi|16303264| VAFSCHCQVQWCCYVECCQCVBELVYTCKH
gi|3915306| VAFSCHCQVQWCCYVECCQCVBELVYTCKH
gi|15082261| VTRFCQCVQWCCYVECCQCTOREEVYTCKG
gi|139748| VTERCHCTFNWCCVVTCLNCTSSQIUECL
gi|3024851| VAFSCHCQV

40

45

50

55

60

65

VEAVMSALECQFQFRFRWNCTLEGRYRASLLKRGFKETAFLYAISSAGLTHALAKACSAGRMRCTCDE
APDLENREGWKWGGCSEIEFGGMVSREFADARENRPDARSAMNRHNEAGRQVIKAGVETTCCKHGVSGS
CTVRTCWRQLAPFHEVGKHLKHYESALKVGGSTTNEAAGEAGAISSPPRGRASGAGSDPLPRTPELVHLLD
SPSFCLAGRFSPGTAGRRCHREKNCEISCCGRGHNTQSRVVTRPCQCQVRWCCYVECRQCTQREEVYTCKG

The disclosed NOV8 amino acid sequence has 270 of 354 amino acid residues (76%) identical to, and 310 of 354 amino acid residues (87%) similar to, the 354 amino acid residue ptnr:SWISSPROT-ACC:O42280 protein from *Gallus gallus* (Chicken) (WNT-14 Protein Precursor (1.2e⁻¹⁵¹).

NOV8 is expressed in at least brain. This information was derived by determining the tissue sources of the sequences that were included in the invention including but not limited to SeqCalling sources, Public EST sources, Literature sources, and/or RACE sources.

In addition, the sequence is predicted to be expressed in brain because of the expression pattern of (GENBANK-ID: gb:GENBANK-ID:AF031168|acc:AF031168.1) a closely related [*Gallus gallus* Wnt-14 protein (Wnt-14) mRNA, complete cds].

NOV8 also has homology to the amino acid sequence shown in the BLASTP data listed in Table 8C.

Table 8C. BLAST results for NOV8					
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi 15082261 ref NP_003386.1 (NM_003395)	wingless-type MMTV integration site family, member 14 [Homo sapiens]	365	306/340 (90%)	321/340 (94%)	e-167
gi 3915306 sp O42280 WN14_CHICK	WNT-14 PROTEIN PRECURSOR	354	270/357 (75%)	310/357 (86%)	e-142
gi 16303264 dbj BAB70499.1 (AB063483)	WNT14B [Homo sapiens]	357	193/339 (56%)	244/339 (71%)	e-100
gi 7106447 ref NP_033548.1 (NM_009522)	wingless-related MMTV integration site 3A [Mus musculus]	352	141/311 (45%)	179/311 (57%)	2e-62
gi 5821261 dbj BAA83743.1 (AB024080)	Wnt-3a [Gallus gallus]	376	139/311 (44%)	179/311 (56%)	3e-62

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 8D.

Table 8D. Information for the ClustalW proteins

- 1) NOV8 (SEQ ID NO:22)
- 2) gi|15082261|ref|NP_003386.1| (NM_003395) wingless-type MMTV integration site family, member 14 [Homo sapiens] (SEQ ID NO:64)
- 3) gi|3915306|sp|O42280|WN14_CHICK WNT-14 PROTEIN PRECURSOR (SEQ ID NO:63)
- 4) gi|16303264|dbj|BAB70499.1| (AB063483) WNT14B [Homo sapiens] (SEQ ID NO:62)
- 5) gi|7106447|ref|NP_033548.1| (NM_009522) wingless-related MMTV integration site 3A [Mus musculus] (SEQ ID NO:67)
- 6) gi|5821261|dbj|BAA83743.1| (AB024080) Wnt-3a [Gallus gallus] (SEQ ID NO:68)

[illegible]

	P25098	QKLGYLFRDFCLNHLEEARPLVEFYEEIKKYEKLETEEERVARSRREIFDSYIMKELLAC	120
	Q99LL8	QKLGYLFRDFCLNHLEEARPLVEFYEEIKKYEKLETEEERVARSRREIFDSYIMKELLAC	118
	P26817	QKLGYLFRDFCLNHLEEARPLVEFYEEIKKYEKLETEEERVARSRREIFDSYIMKELLAC	120
5	NOV9	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQNLRGDVFOKFIESDKFTRFCQWKNV	180
	A53791	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQNLRGDVFOKFIESDKFTRFCQWKNV	180
	P25098	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQNLRGDVFOKFIESDKFTRFCQWKNV	180
	Q99LL8	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQNLRGDVFOKFIESDKFTRFCQWKNV	178
	P26817	SHPFKSATEHVQGHGKQVPPDLFQPYIEEICQNLRGDVFOKFIESDKFTRFCQWKNV	180
10	NOV9	ELNIHLTMNDFSVHRIIGRGGFGEVYGCRAKDTGKMYAMKCLDKKRIKMKQGETLALNER	240
	A53791	ELNIHLTMNDFSVHRIIGRGGFGEVYGCRAKDTGKMYAMKCLDKKRIKMKQGETLALNER	240
	P25098	ELNIHLTMNDFSVHRIIGRGGFGEVYGCRAKDTGKMYAMKCLDKKRIKMKQGETLALNER	240
	Q99LL8	ELNIHLTMNDFSVHRIIGRGGFGEVYGCRAKDTGKMYAMKCLDKKRIKMKQGETLALNER	238
15	P26817	ELNIHLTMNDFSVHRIIGRGGFGEVYGCRAKDTGKMYAMKCLDKKRIKMKQGETLALNER	240
	NOV9	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
	A53791	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
	P25098	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
20	Q99LL8	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	298
	P26817	IMLSLVSTGDCPFIVCMSYAFHTPDKLSFILDLMNGGDLHYHLSQHGVFSEADMRFYAAE	300
	NOV9	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRIIDLGLACDFSKKKPHASVGTGHYMAPE	360
	A53791	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRIIDLGLACDFSKKKPHASVGTGHYMAPE	360
25	P25098	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRIIDLGLACDFSKKKPHASVGTGHYMAPE	360
	Q99LL8	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRIIDLGLACDFSKKKPHASVGTGHYMAPE	358
	P26817	IILGLEHMHNRFFVYRDLKPANILLDEHGHVRIIDLGLACDFSKKKPHASVGTGHYMAPE	360
	NOV9	VLQKGVA YDSSADWFLGCMFLKLLRGHSPFRQHKTKDKHEIDRMTLTMAVELPDSFSPE	420
30	A53791	VLQKGVA YDSSADWFLGCMFLKLLRGHSPFRQHKTKDKHEIDRMTLTMAVELPDSFSPE	420
	P25098	VLQKGVA YDSSADWFLGCMFLKLLRGHSPFRQHKTKDKHEIDRMTLTMAVELPDSFSPE	420
	Q99LL8	VLQKGVA YDSSADWFLGCMFLKLLRGHSPFRQHKTKDKHEIDRMTLTMAVELPDSFSPE	418
	P26817	VLQKGVA YDSSADWFLGCMFLKLLRGHSPFRQHKTKDKHEIDRMTLTMAVELPDSFSPE	420
35	NOV9	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKESPFERSLDWQMVFLQKYPPLIPPRGEVNAA	480
	A53791	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKESPFERSLDWQMVFLQKYPPLIPPRGEVNAA	480
	P25098	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKESPFERSLDWQMVFLQKYPPLIPPRGEVNAA	480
	Q99LL8	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKESPFERSLDWQMVFLQKYPPLIPPRGEVNAA	478
	P26817	LRSLLLEGLLQRDVNRRLGCLGRGAQEVKESPFERSLDWQMVFLQKYPPLIPPRGEVNAA	480
40	NOV9	DAFDIGSFDEEDTKGIR-----QEAETVFDTINAETDRLEARK	519
	A53791	DAFDIGSFDEEDTKGIR-----QEAETVFDTINAETDRLEARK	540
	P25098	DAFDIGSFDEEDTKGIR-----QEAETVFDTINAETDRLEARK	540
45	Q99LL8	DAFDIGSFDEEDTKGIR-----QEAETVFDTINAETDRLEARK	538
	P26817	DAFDIGSFDEEDTKGIR-----QEAETVFDTINAETDRLEARK	540
	NOV9	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	579
	A53791	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	600
	P25098	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	600
50	Q99LL8	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	598
	P26817	KAKNKQLGHEEDYALGKDCIMHGYSKMGNPFLTQWQRRYFYLFENRLEWRGEGEAPQSL	600
	NOV9	LTMEEQSVETQIKERKCLLLKIRGGKQFVLQCDSDPELVQWKKELRDAYREAQQLVQR	639
	A53791	LTMEEQSVETQIKERKCLLLKIRGGKQFVLQCDSDPELVQWKKELRDAYREAQQLVQR	660
55	P25098	LTMEEQSVETQIKERKCLLLKIRGGKQFVLQCDSDPELVQWKKELRDAYREAQQLVQR	660
	Q99LL8	LTMEEQSVETQIKERKCLLLKIRGGKQFVLQCDSDPELVQWKKELRDAYREAQQLVQR	658
	P26817	LTMEEQSVETQIKERKCLLLKIRGGKQFVLQCDSDPELVQWKKELRDAYREAQQLVQR	660
	NOV9	VPKMKNKPRSPVVELSKVPLVQRGSANGL	668
60	A53791	VPKMKNKPRSPVVELSKVPLVQRGSANGL	689
	P25098	VPKMKNKPRSPVVELSKVPLVQRGSANGL	689
	Q99LL8	VPKMKNKPRSPVVELSKVPLVQRGSANGL	687
	P26817	VPKMKNKPRSPVVELSKVPLVQRGSANGL	689

65 Tables 9E-9L list the domain descriptions from DOMAIN analysis results against NOV9. This indicates that the NOV9 sequence has properties similar to those of other proteins known to contain this domain.

Table 10C. BLAST results for NOV10

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
ptnr:SP TREMBL-ACC:Q9UL64	ALPHA-MANNOSIDASE 6A8B - Homo sapiens	1062	763/771 (99%)	767/771 (99%)	0.0
ptnr:SP TREMBL-ACC:Q9NTJ4	HYPOTHETICAL 115.8 KDA PROTEIN - Homo sapiens	1040	715/722 (99%)	718/722 (99%)	0.0
ptnr:TREMBLNEW-ACC:AAH16253	SIMILAR TO MANNOSIDASE, ALPHA, CLASS 2C, MEMBER 1	1039	635/730 (89%)	692/730 (94%)	0.0
ptnr:SWISSPROT-ACC:P21139	Alpha-mannosidase (EC 3.2.1.24)	1040	625/731 (85%)	661/731 (90%)	0.0
ptnr:SP TREMBL-ACC:Q13358	ALPHA-MANNOSIDASE - Homo sapiens	425	425/425 (100%)	425/425 (100%)	0.0

The homology between these and other sequences is shown graphically in the ClustalW analysis shown in Table 10D. In the ClustalW alignment of the NOV10 protein, as well as all other ClustalW analyses herein, the black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be altered to a much broader extent without altering protein structure or function.

Table 10D. ClustalW Analysis of NOV10

- 1) NOV10 (SEQ ID NO:26)
- 2) ptnr: ALPHA-MANNOSIDASE 6A8B - Homo sapiens (SEQ ID NO:73)
- 3) ptnr: HYPOTHETICAL 115.8 KDA PROTEIN - Homo sapiens (SEQ ID NO:74)
- 4) ptnr: SIMILAR TO MANNOSIDASE, ALPHA, CLASS 2C, MEMBER 1 (SEQ ID NO:75)
- 5) ptnr: Alpha-mannosidase (EC 3.2.1.24) (SEQ ID NO:76)

NOV10	MAAAPFLKHH	RTTTERVEKF	VSPIYFTDN	LRGRFCASCPVAVLSSFLT	ERLPYCEAVQR	-----	FRPAQVG	70
Q9UL64	MAAAPFLKHH	RTTTERVEKF	VSPIYFTDN	LRGRFCASCPVAVLSSFLT	ERLPYCEAVQR	-----	FRPAQVG	70
Q9NTJ4	MAAAPFLKHH	RTTTERVEKF	VSPIYFTDN	LRGRFCASCPVAVLSSFLT	ERLPYCEAVQR	-----	FRPAQVG	70
AAH16253	MAAAPFLKHH	RTTTERVEKF	VSPIYFTDN	LRGRFCASCPVAVLSSFLT	ERLPYCEAVQR	-----	FRPAQVG	69
P21139	MAAAPFLKHH	RTTTERVEKF	VSPIYFTDN	LRGRFCASCPVAVLSSFLT	ERLPYCEAVQR	-----	FRPAQVG	79
NOV10	D-SFGPTWWTCT	FRVELTIP	PAWVQGEVHLCWE	SDGEGLVWRE	GEVQGLTH	-----	EGEKTSYVL	140
Q9UL64	D-SFGPTWWTCT	FRVELTIP	PAWVQGEVHLCWE	SDGEGLVWRE	GEVQGLTH	-----	EGEKTSYVL	140
Q9NTJ4	D-SFGPTWWTCT	FRVELTIP	PAWVQGEVHLCWE	SDGEGLVWRE	GEVQGLTH	-----	EGEKTSYVL	140
AAH16253	D-SFGPTWWTCT	FRVELTIP	PAWVQGEVHLCWE	SDGEGLVWRE	GEVQGLTH	-----	EGEKTSYVL	139
P21139	D-SFGPTWWTCT	FRVELTIP	PAWVQGEVHLCWE	SDGEGLVWRE	GEVQGLTH	-----	EGEKTSYVL	159
NOV10	S-LTLYVEVAGNLLGAGKGS	IAAPDPERMF	QLSRAELAVF	-----	HRDVHMLLVE	LELLLGIAF	-----	200
Q9UL64	S-LTLYVEVAGNLLGAGKGS	IAAPDPERMF	QLSRAELAVF	-----	HRDVHMLLVE	LELLLGIAKCLQ	-----	211
Q9NTJ4	S-LTLYVEVAGNLLGAGKGS	IAAPDPERMF	QLSRAELAVF	-----	HRDVHMLLVE	LELLLGIAKGLG	-----	211
AAH16253	S-LTLYVEVAGNLLGAGKGS	IAAPDPERMF	QLSRAELAVF	-----	HRDVHMLLVE	LELLLGIAKGLG	-----	210
P21139	S-LTLYVEVAGNLLGAGKGS	IAAPDPERMF	QLSRAELAVF	-----	HRDVHMLLVE	LELLLGIAKGLG	-----	229
NOV10	ALYTANOMVWCDPAQPET	IPVACALAS	FFGCHGGSQHT	IHAQCHCH	IDTAWLWPF	KETVRKCARSW	VFALQVLRNP	291
Q9UL64	ALYTANOMVWCDPAQPET	IPVACALAS	FFGCHGGSQHT	IHAQCHCH	IDTAWLWPF	KETVRKCARSW	VFALQVLRNP	291
Q9NTJ4	ALYTANOMVWCDPAQPET	IPVACALAS	FFGCHGGSQHT	IHAQCHCH	IDTAWLWPF	KETVRKCARSW	VFALQVLRNP	291
AAH16253	ALYTANOMVWCDPAQPET	IPVACALAS	FFGCHGGSQHT	IHAQCHCH	IDTAWLWPF	KETVRKCARSW	VFALQVLRNP	290
P21139	ALYTANOMVWCDPAQPET	IPVACALAS	FFGCHGGSQHT	IHAQCHCH	IDTAWLWPF	KETVRKCARSW	VFALQVLRNP	229
NOV10	-----	AQQLRW/KSRYPGLYSR	IQEFACRQGFVPYGGT	WVENDQ	LPSGEAMVRQFLQGN	FFLQEFGRACSEFWLF	-----	272
Q9UL64	-----	EFIFACSNQOLEWV/KSRYPGLYSR	IQEFACRQGFVPYGGT	WVENDQ	LPSGEAMVRQFLQGN	FFLQEFGRACSEFWLF	-----	371

5	Q9NTJ4	SFIACEQAQQLLEWVSKRYPGLYSRIOEFACRGQDFVPGGTVNDONLPSGEAMVRFQLOGNFFLOEFCKMCSFFNLF	371
	AAH16253	SFIACEQAQQLLEWVSKRYPGLYSRIOEFACRGQDFVPGGTVNDONLPSGEAMVRFQLOGNFFLOEFCKMCSFFNLF	370
	P21139	-----TLGGDSSPPQALLT-----ANDVNDGDP-AQSPSTFPAASFTAB-----RFGDQRO-----	281
10	NOV10	DTFGYSAQLPQIMHCCGILRRFLTQKLSNVLVNSFPHTTFFWEGLDGSRLVLFHPPGDSYCMGCSVEEVLKTVNNRDKGF	352
	Q9UL64	DTFGYSAQLPQIMHCCGILRRFLTQKLSNVLVNSFPHTTFFWEGLDGSRLVLFHPPGDSYCMGCSVEEVLKTVNNRDKGF	451
	Q9NTJ4	DTFGYSAQLPQIMHCCGILRRFLTQKLSNVLVNSFPHTTFFWEGLDGSRLVLFHPPGDSYCMGCSVEEVLKTVNNRDKGF	451
15	AAH16253	DTFGYSAQLPQIMHCCGILRRFLTQKLSNVLVNSFPHTTFFWEGLDGSRLVLFHPPGDSYCMGCSVEEVLKTVNNRDKGF	450
	P21139	-----GSSQHTLILATGCHILDTATL-----P-----FKEITVAKCAR-----	314
20	NOV10	AMHSAFLEFGFGDGGCGPTQTMLDRLKRLSNTDGLPRVOLSSSPRCFLSALESDSQELCTHVGELFLLEHNGTYYTHAQIKH	432
	Q9UL64	AMHSAFLEFGFGDGGCGPTQTMLDRLKRLSNTDGLPRVOLSSSPRCFLSALESDSQELCTHVGELFLLEHNGTYYTHAQIKH	531
	Q9NTJ4	AMHSAFLEFGFGDGGCGPTQTMLDRLKRLSNTDGLPRVOLSSSPRCFLSALESDSQELCTHVGELFLLEHNGTYYTHAQIKH	531
25	AAH16253	AMHSAFLEFGFGDGGCGPTQTMLDRLKRLSNTDGLPRVOLSSSPRCFLSALESDSQELCTHVGELFLLEHNGTYYTHAQIKH	530
	P21139	-----TNISGFLFGFGDGGCGPTQTMLDRLKRLSNTDGLPRVOLSSSPRCFLSALESDSQELCTHVGELFLLEHNGTYYTHAQIKH-----	325
30	NOV10	ENRECEIILHDVBLSSLLALARSQFLYPAAQLOHLWRLLLLNFHDVVTGSCIQMWAEARACHYEDIRSHQNTLLSAAH	512
	Q9UL64	ENRECEIILHDVBLSSLLALARSQFLYPAAQLOHLWRLLLLNFHDVVTGSCIQMWAEARACHYEDIRSHQNTLLSAAH	611
	Q9NTJ4	ENRECEIILHDVBLSSLLALARSQFLYPAAQLOHLWRLLLLNFHDVVTGSCIQMWAEARACHYEDIRSHQNTLLSAAH	611
35	AAH16253	ENRECEIILHDVBLSSLLALARSQFLYPAAQLOHLWRLLLLNFHDVVTGSCIQMWAEARACHYEDIRSHQNTLLSAAH	610
	P21139	-----RN-----TEPTFAECBOGATB-----ETDLSMK-----NGP-----EG-----LVAQZGFAN-----	365
40	NOV10	AALCAGEPGPGGLLIIVNTPWKRIEYMALPKPGGAHSLALVTPSGMAYAPVPPPTSLQPLLPODPVFVVOETGGSVTLCN	592
	Q9UL64	AALCAGEPGPGGLLIIVNTPWKRIEYMALPKPGGAHSLALVTPSGMAYAPVPPPTSLQPLLPODPVFVVOETGGSVTLCN	691
	Q9NTJ4	AALCAGEPGPGGLLIIVNTPWKRIEYMALPKPGGAHSLALVTPSGMAYAPVPPPTSLQPLLPODPVFVVOETGGSVTLCN	691
45	AAH16253	AALCAGEPGPGGLLIIVNTPWKRIEYMALPKPGGAHSLALVTPSGMAYAPVPPPTSLQPLLPODPVFVVOETGGSVTLCN	690
	P21139	-----AALCAGEPGPGGLLIIVNTPWKRIEYMALPKPGGAHSLALVTPSGMAYAPVPPPTSLQPLLPODPVFVVOETGGSVTLCN-----	374
50	NOV10	GIIIRVKLDPTGRLTSLVLVASGREAI AEGAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGGLRGSA	772
	Q9UL64	GIIIRVKLDPTGRLTSLVLVASGREAI AEGAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGGLRGSA	671
	Q9NTJ4	GIIIRVKLDPTGRLTSLVLVASGREAI AEGAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGGLRGSA	771
55	AAH16253	GIIIRVKLDPTGRLTSLVLVASGREAI AEGAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGGLRGSA	770
	P21139	-----GIIIRVKLDPTGRLTSLVLVASGREAI AEGAVGNQFVLFDDVPLYWDADVMDYHLETRKPVLGQAGTLAVGTGGLRGSA-----	402
60	NOV10	WFLLOIISNRSRLSQEVVLVDGCPYVRFHTEVHHWEAHKFLKVEFPARVRSQATYIICFGHQLRPTHYNTSDHARFEVH	752
	Q9UL64	WFLLOIISNRSRLSQEVVLVDGCPYVRFHTEVHHWEAHKFLKVEFPARVRSQATYIICFGHQLRPTHYNTSDHARFEVH	851
	Q9NTJ4	WFLLOIISNRSRLSQEVVLVDGCPYVRFHTEVHHWEAHKFLKVEFPARVRSQATYIICFGHQLRPTHYNTSDHARFEVH	851
65	AAH16253	WFLLOIISNRSRLSQEVVLVDGCPYVRFHTEVHHWEAHKFLKVEFPARVRSQATYIICFGHQLRPTHYNTSDHARFEVH	850
	P21139	-----WFLLOIISNRSRLSQEVVLVDGCPYVRFHTEVHHWEAHKFLKVEFPARVRSQATYIICFGHQLRPTHYNTSDHARFEVH-----	408
70	NOV10	AHRWMDLSEHGFGCLALLNDCKYCASVRGSLLSLLRAPKAPDATADTGRHEFTYALMPHKGSFODACVIOAAYSUNFPI	832
	Q9UL64	AHRWMDLSEHGFGCLALLNDCKYCASVRGSLLSLLRAPKAPDATADTGRHEFTYALMPHKGSFODACVIOAAYSUNFPI	931
	Q9NTJ4	AHRWMDLSEHGFGCLALLNDCKYCASVRGSLLSLLRAPKAPDATADTGRHEFTYALMPHKGSFODACVIOAAYSUNFPI	931
75	AAH16253	AHRWMDLSEHGFGCLALLNDCKYCASVRGSLLSLLRAPKAPDATADTGRHEFTYALMPHKGSFODACVIOAAYSUNFPI	930
	P21139	-----AHRWMDLSEHGFGCLALLNDCKYCASVRGSLLSLLRAPKAPDATADTGRHEFTYALMPHKGSFODACVIOAAYSUNFPI-----	408
80	NOV10	LALPAPSPAPATSNWAFSVSSPAVVLVTVKQAESSPORSLVLRLYEAHGSHVDGNLHLSLPVQEAITCDLLEPPDACH	912
	Q9UL64	LALPAPSPAPATSNWAFSVSSPAVVLVTVKQAESSPORSLVLRLYEAHGSHVDGNLHLSLPVQEAITCDLLEPPDACH	1011
	Q9NTJ4	LALPAPSPAPATSNWAFSVSSPAVVLVTVKQAESSPORSLVLRLYEAHGSHVDGNLHLSLPVQEAITCDLLEPPDACH	1011
85	AAH16253	LALPAPSPAPATSNWAFSVSSPAVVLVTVKQAESSPORSLVLRLYEAHGSHVDGNLHLSLPVQEAITCDLLEPPDACH	1010
	P21139	-----LALPAPSPAPATSNWAFSVSSPAVVLVTVKQAESSPORSLVLRLYEAHGSHVDGNLHLSLPVQEAITCDLLEPPDACH-----	408
90	NOV10	LISGQPPBAHLPSLPSAVPPIARASASATLISYWGNGFVCRRLNGLLISASFA- 963	
	Q9UL64	LISGQPPBAHLPSLPSAVPPIARASASATLISYWGNGFVCRRLNGLLISASFA- 1062	
	Q9NTJ4	-----LDNRRLKLLDFSP-----QVLS-----LILYLQSPH 1040	
95	AAH16253	-----LDNRRLKLLDFSP-----QVLS-----LILYLQSPH 1039	
	P21139	-----LDNRRLKLLDFSP-----QVLS-----LILYLQSPH 408	

60 Table 10E lists the domain description from DOMAIN analysis results against NOV10. This indicates that the NOV10 sequence has properties similar to those of other proteins known to contain this domain.

Table 10E. Domain Analysis of NOV10				
Model	Description	Score	E-value	
Glyco_hydro_38 (InterPro)	Glycosyl hydrolases family 38	140.5	1e-39	
(SEQ ID NO:111)				
Glyco_hydro_38: domain 1 of 2, from 230 to 332: score 89.2, E = 5.4e-25				
	*->vtGGVVMNDEAttHyedlIdQlIteGHqfLeenfGsdvKPkvgWsIDP + + + + +++ ++++ ++ + + ++ + + ++ +			
AC058790_d	230 VGGTWVEMDGNLPSGEAMVRQFLQGQNFLLQEGF--XMCSEFWLPDT	274		
	FGHSatmPyLlraqaGfdgflIqRihYadKksfaetkqleFvWRqswslt			
	+ ++ ++ + + + + +++++ + ++++ + +			
AC058790_d	275 FGYSAPQLPQIM-HGCGIRRFLLTQKLSWNVLNVSFPHHT---PFWF---GLD	317		



Q9ZSN4 AIAQADQNYDYASNSVILHLDAGDEVFIKLDGGKHHGGNSNKYSTFSGFIIYSD--- 258
P02746 TFCDYAYNTEQVTGGMVLKLEGGENVFLCATDKNSLLGMEGANSIFSGFLLFPDMEA 251
AAH08983 TFCDYAYNTEQVTGGMVLKLEGGENVFLCATDKNSLLGMEGANSIFSGFLLFPDMEA 253

5

Tables 11E-11F list the domain descriptions from DOMAIN analysis results against NOV11. This indicates that the NOV11 sequence has properties similar to those of other proteins known to contain this domain.

Table 11E. Domain Analysis of NOV11

gnl|Smart|smart00110, Clq, Complement component Clq domain.; Globular domain found in many collagens and eponymously in complement Clq. When part of full length proteins these domains form a 'bouquet' due to the multimerization of heterotrimers. The Clq fold is similar to that of tumour necrosis factor. (SEQ ID NO:104)
CD-Length = 132 residues, 99.2% aligned
Score = 113 bits (283), Expect = 1e-26

Query: 108 PRIAFYAGL--RRPHEGYEVLRFDDVVTNVGNAYEAASGKFTCPMPGVYFFAYHVLMRGG 165
|| || || + + ||| + | | + + ||||| + ||| + | + | +
15 Sbjct: 2 PRSAPSVIRSTNRPPPGQPVRFDKVLNQGHYDPSTGKFTCPVPGVYFFSYHIESK-- 59
Query: 166 DGTSMWADLMKNGQVRASAIQDADQNYDYASNSVILHLDVGDEVFIKLDGGKVHG--GNT 224
| ++ ||||| + | || + | ||| + + + || |
20 Sbjct: 60 -GRNVKVSMLKNGIQVMRECDEYQKGLYQVASGGALLQLRQGDQVWLELDDKKNGLYAGE 118
Query: 225 NKYSTFSGFIIYPD 238
||||| + + + ||
Sbjct: 119 EVDSTFSGFLLFPD 132

25

Table 11F. Domain Analysis of NOV11

gnl|Pfam|pfam00386, Clq, Clq domain. Clq is a subunit of the C1 enzyme complex that activates the serum complement system. (SEQ ID NO:112)
CD-Length = 125 residues, 100.0% aligned
Score = 102 bits (253), Expect = 3e-23

Query: 111 AFYAGLR-RPHEGYEVLRFDDVVTNVGNAYEAASGKFTCPMPGVYFFAYHVLMRGGDGTS 169
|| | || + + ||| + | | + + ||||| + ||| + | + | + || +
30 Sbjct: 1 AFTAIRSTRPPAPGQPVIFDEVLYNQGHYDPATGKFTCPVPGLYYFNHVSCK--GTN 57
Query: 170 MWADLMKNGQVRASAIQDADQNYDYASNSVILHLDVGDEVFIKLDGGKVHG--GNTNKY 227
+ ||| + | + | | || + | ||| + + + || + + | +
Sbjct: 58 VCVSLMRNGVPVMSFCDEYAKGTYQVASGGAVLQLRQGDQVWLELDDKQTNGLLGEGVH 117
Query: 228 STFSGFII 235
| |||| + +
35 Sbjct: 118 SVFSGFLL 125

The first component of complement system is a calcium-dependent complex of the 3 subcomponents Clq,Clr, and C1s. Subcomponent Clq binds to immunoglobulin complexes

ptnr: SPTREMBL- ACC: P70208	PLEXIN 3 - Mus musculus	1872	1245/1874 (66%)	1478/1874 (78%)	0.0
--------------------------------	----------------------------	------	--------------------	--------------------	-----

The homology between these and other sequences is shown graphically in the ClustalW analysis shown in Table 12D. In the ClustalW alignment of the NOV12 protein, as well as all other ClustalW analyses herein, the black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties), whereas non-highlighted amino acid residues are less conserved and can potentially be altered to a much broader extent without altering protein structure or function.

Table 12D. ClustalW Analysis of NOV12

10	1) NOV12 (SEQ ID NO:32)
	2) ptnr: PLEXIN 1 - Mus musculus (SEQ ID NO:81)
	3) ptnr: NOV/PLEXIN-A1 PROTEIN - Homo sapiens (SEQ ID NO:82)
	4) ptnr: PLEXIN PRECURSOR - Xenopus laevis (SEQ ID NO:83)
15	5) ptnr: Plexin A3 precursor (Plexin 4) (SEQ ID NO:84)
6)	
	NOV12 NMLTPAQPEHRGPRPQAMPDPPPSLQVLLLLLLLLTTPQKAAEALPRACGCSOPPFRTFSASDWGLTHLVVHEOTGR 60
20	P70206 -----PLPPLSSRTLLELLLTTRVMIRITSSPPAGLGPQAFRTFVASDWGLTHLVVHEOTGR 60
	Q9UIW2 -----LAAEALPRACGCSOPPFRTFASDWGLTHLVVHEOTGR 40
	Q91823 -----LLHAERPLPPIITFTLVLTGSSNITP-----SDGSPKDFRTFASDWGLTHLVVHEOTGR 55
	P51805 -----PSQCLLLPLPLAVE-----EALGNRPFAFVVEETLTHLVVHEOTGR 44
	NOV12 VYVGAVNR IYKLSCHLTLLRAHVTPGVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 160
25	P70206 VYVGAVNR IYKLSCHLTLLRAHVTPGVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 140
	Q9UIW2 VYVGAVNR IYKLSCHLTLLRAHVTPGVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 120
	Q91823 VYVGAVNR IYKLSCHLTLLRAHVTPGVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 135
	P51805 VYVGAVNR IYKLSCHLTLLRAHVTPGVEDNEKCYPPPSVOSCPHGLCSTDNVKKLLLDYAAARLLACGSASQGI COFLR 124
30	NOV12 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMGFVYQ 240
	P70206 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMGFVYQ 220
	Q9UIW2 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMGFVYQ 200
	Q91823 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMGFVYQ 160
35	P51805 LDDLFLKLGEPHHRKEHYLSSVQEGAGSMAGVLIAGPGGQAKLFVGTPI DGKSEYFPTLSSRRLLMANEEDACMGFVYQ 202
	NOV12 EFVSSOLKI PSOTLSKFPAPDIYVYVSFRSEQFVYVLTQLDTOLTSPOAAGEHFTSKIIVRLCVDOPKFYSYVEFPICG 320
	P70206 EFVSSOLKI PSOTLSKFPAPDIYVYVSFRSEQFVYVLTQLDTOLTSPOAAGEHFTSKIIVRLCVDOPKFYSYVEFPICG 300
	Q9UIW2 EFVSSOLKI PSOTLSKFPAPDIYVYVSFRSEQFVYVLTQLDTOLTSPOAAGEHFTSKIIVRLCVDOPKFYSYVEFPICG 280
	Q91823 EFVSSOLKI PSOTLSKFPAPDIYVYVSFRSEQFVYVLTQLDTOLTSPOAAGEHFTSKIIVRLCVDOPKFYSYVEFPICG 160
40	P51805 EFVSSOLKI PSOTLSKFPAPDIYVYVSFRSEQFVYVLTQLDTOLTSPOAAGEHFTSKIIVRLCVDOPKFYSYVEFPICG 282
	NOV12 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOCCNRVKPPKESALCLFLTRAIKEIKERIOSCYRGECH 400
	P70206 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOCCNRVKPPKESALCLFLTRAIKEIKERIOSCYRGECH 380
	Q9UIW2 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOCCNRVKPPKESALCLFLTRAIKEIKERIOSCYRGECH 360
45	Q91823 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOCCNRVKPPKESALCLFLTRAIKEIKERIOSCYRGECH 160
	P51805 EOAGVEYRLVQDAYLSRPGCALAHOLGLAEDVDLFTVFAOCCNRVKPPKESALCLFLTRAIKEIKERIOSCYRGECH 362
	NOV12 LSLPWLINKELGGINSPLOIDDDFCQDQNPOLGGTVTIECTPLFVDNDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 480
	P70206 LSLPWLINKELGGINSPLOIDDDFCQDQNPOLGGTVTIECTPLFVDNDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 460
50	Q9UIW2 LSLPWLINKELGGINSPLOIDDDFCQDQNPOLGGTVTIECTPLFVDNDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 440
	Q91823 LSLPWLINKELGGINSPLOIDDDFCQDQNPOLGGTVTIECTPLFVDNDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 160
	P51805 LSLPWLINKELGGINSPLOIDDDFCQDQNPOLGGTVTIECTPLFVDNDGLTAVAAAYDYGRTVVFACTRSGRIRKILV 442
	NOV12 DLSNIEGRPALAYESVVAOEGSPIRLDVLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 560
55	P70206 DLSNIEGRPALAYESVVAOEGSPIRLDVLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 540
	Q9UIW2 DLSNIEGRPALAYESVVAOEGSPIRLDVLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 520
	Q91823 DLSNIEGRPALAYESVVAOEGSPIRLDVLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 160
	P51805 DLSNIEGRPALAYESVVAOEGSPIRLDVLVLSNHOVLYAXTEKQVTVPVVESCVOYTSCELCIGSRDPHCGNCVLIHSICS 518
60	NOV12 RRDACERAEPPORFASDLIQCQVLTQVPPNVSVTMSQVFVLVLQANVPDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 639
	P70206 RRDACERAEPPORFASDLIQCQVLTQVPPNVSVTMSQVFVLVLQANVPDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 618
	Q9UIW2 RRDACERAEPPORFASDLIQCQVLTQVPPNVSVTMSQVFVLVLQANVPDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 598
	Q91823 RRDACERAEPPORFASDLIQCQVLTQVPPNVSVTMSQVFVLVLQANVPDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 160
65	P51805 RRDACERAEPPORFASDLIQCQVLTQVPPNVSVTMSQVFVLVLQANVPDLSAGVNCSEFDFTESESVLE-DGRIHQHRSF 597
	NOV12 SAREVAPITRGG--SDQRVVKLYLKSKEKGKFFASVDVFVYVCSVHOS--GLACVNGSFPCHWCKYRHVCTNMADCAFLR 719
	P70206 SAREVAPITRGG--SDQRVVKLYLKSKEKGKFFASVDVFVYVCSVHOS--GLACVNGSFPCHWCKYRHVCTNMADCAFLR 695

Q9UIW2	SAREVAPITRGC--GDORVVKLYNKSKEGKGFASVDFVFNCSVHCS	CLSCVNGSPFCHCKYRHVC	DNVADCAFL	675
Q91823	-----	-----	-----	160
P51805	SLGGRAITRCH--GATETVNLQSLKSTGVRFADDFVFNCSVLAG	SYGCVOSPPCHCKYRHVC	SRPHCESSQ	674
NOV12	GRVWSEDCPOILPSTQIYVPVGVVKPITLAAPNLPOPOSGQGYECLFHI	PGSPARVTALRFN	SSSLOCCSSSYSEGN	799
P70206	GRVWSEDCPOILPSTQIYVPVGVVKPITLAAPNLPOPOSGQGYECLFHI	PGSPARVTALRFN	SSSLOCCSSSYSEGN	775
Q9UIW2	GRVWSEDCPOILPSTQIYVPVGVVKPITLAAPNLPOPOSGQGYECLFHI	PGSPARVTALRFN	SSSLOCCSSSYSEGN	755
Q91823	-----	-----	-----	160
P51805	GRVWSEDCPOILPSTQIYVPVGVVKPITLAAPNLPOPOSGQGYECLFHI	PGSPARVTALRFN	SSSLOCCSSSYSEGN	754
NOV12	DVSDLPVNLVVVWGNFVIDNPQIOAHLYKCPALPSSCGCLCKADPRFECCW	CVARRCSLRHHIC	ADTPASMKHAKIC	879
P70206	DVSDLPVNLVVVWGNFVIDNPQIOAHLYKCPALPSSCGCLCKADPRFECCW	CVARRCSLRHHIC	ADTPASMKHAKIC	855
Q9UIW2	DVSDLPVNLVVVWGNFVIDNPQIOAHLYKCPALPSSCGCLCKADPRFECCW	CVARRCSLRHHIC	ADTPASMKHAKIC	835
Q91823	-----	-----	-----	160
P51805	HHGCTEDFVVVWGNFVIDNPQIOAHLYKCPALPSSCGCLCKADPRFECCW	CVARRCSLRHHIC	ADTPASMKHAKIC	833
NOV12	SSRCTDPKILKLSPETGPROGCTRLTITCEMLGLRPFEDVRLGVVGVKVLCS	PVSESYISAEQIVCEIGDASSVRAHDALV		959
P70206	SSRCTDPKILKLSPETGPROGCTRLTITCEMLGLRPFEDVRLGVVGVKVLCS	PVSESYISAEQIVCEIGDASSVRAHDALV		935
Q9UIW2	SSRCTDPKILKLSPETGPROGCTRLTITCEMLGLRPFEDVRLGVVGVKVLCS	PVSESYISAEQIVCEIGDASSVRAHDALV		915
Q91823	-----	-----	-----	160
P51805	CHGCTEDFVVVWGNFVIDNPQIOAHLYKCPALPSSCGCLCKADPRFECCW	CVARRCSLRHHIC	ADTPASMKHAKIC	911
NOV12	EVCVADCSFHYRALS PKRFTFTVTPFYRVSPSRGPLSGGTWIGIEGSHINAG	SOVAVSVGGRPCSF	SKSRPNSREIRCLT	1039
P70206	EVCVADCSFHYRALS PKRFTFTVTPFYRVSPSRGPLSGGTWIGIEGSHINAG	SOVAVSVGGRPCSF	SKSRPNSREIRCLT	1013
Q9UIW2	EVCVADCSFHYRALS PKRFTFTVTPFYRVSPSRGPLSGGTWIGIEGSHINAG	SOVAVSVGGRPCSF	SKSRPNSREIRCLT	993
Q91823	-----	-----	-----	160
P51805	ELCVADCSFHYRALS PKRFTFTVTPFYRVSPSRGPLSGGTWIGIEGSHINAG	SOVAVSVGGRPCSF	SKSRPNSREIRCLT	989
NOV12	PPCQ--PGSAPITILINRAOITLPEVKYNYTEDPTILRIDPEWSINSGGTLL	TVTCTINLATVREPRI	IRAKYGGIEREN	1117
P70206	PPCQ--PGSAPITILINRAOITLPEVKYNYTEDPTILRIDPEWSINSGGTLL	TVTCTINLATVREPRI	IRAKYGGIEREN	1092
Q9UIW2	PPCQ--PGSAPITILINRAOITLPEVKYNYTEDPTILRIDPEWSINSGGTLL	TVTCTINLATVREPRI	IRAKYGGIEREN	1072
Q91823	-----	-----	-----	160
P51805	PLSTLGSAPITILINRAOITLPEVKYNYTEDPTILRIDPEWSINSGGTLL	TVTCTINLATVREPRI	IRAKYGGIEREN	1069
NOV12	LYNDITMVCRAPSVANDRSPPELGERPDELGFVMDVBSLLVLSISFLYY	PCPVLEPLSPITGLLEKPSPLILKCH		1197
P70206	LYNDITMVCRAPSVANDRSPPELGERPDELGFVMDVBSLLVLSISFLYY	PCPVLEPLSPITGLLEKPSPLILKCH		1172
Q9UIW2	LYNDITMVCRAPSVANDRSPPELGERPDELGFVMDVBSLLVLSISFLYY	PCPVLEPLSPITGLLEKPSPLILKCH		1152
Q91823	-----	-----	-----	160
P51805	QINDITMVCRAPSVANDRSPPELGERPDELGFVMDVBSLLVLSISFLYY	PCPVLEPLSPITGLLEKPSPLILKCH		1149
NOV12	NLLPPAPONSRIYTVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGF	EFSPGTLQVYSDSLTLPA	IVGIGGGG	1277
P70206	NLLPPAPONSRIYTVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGF	EFSPGTLQVYSDSLTLPA	IVGIGGGG	1252
Q9UIW2	NLLPPAPONSRIYTVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGF	EFSPGTLQVYSDSLTLPA	IVGIGGGG	1232
Q91823	-----	-----	-----	160
P51805	NLLPPAPONSRIYTVLIGSTPCITLVSETOLLCEAPNLTGQHKVTVRAGGF	EFSPGTLQVYSDSLTLPA	IVGIGGGG	1229
NOV12	LLLLVIVAVLIAYKKRSRDAORTLKRLOLOMDLESRVALECKEAFASLOT	DIHELINOLDGACI	PFLDYRTYAMRVLF	1357
P70206	LLLLVIVAVLIAYKKRSRDAORTLKRLOLOMDLESRVALECKEAFASLOT	DIHELINOLDGACI	PFLDYRTYAMRVLF	1332
Q9UIW2	LLLLVIVAVLIAYKKRSRDAORTLKRLOLOMDLESRVALECKEAFASLOT	DIHELINOLDGACI	PFLDYRTYAMRVLF	1312
Q91823	-----	-----	-----	160
P51805	LLLLVIVAVLIAYKKRSRDAORTLKRLOLOMDLESRVALECKEAFASLOT	DIHELINOLDGACI	PFLDYRTYAMRVLF	1309
NOV12	GLEDHPVLKEMEVOANVEKSLTLFQOLLTKKHFLITRTLEAORSFSMRDRG	VASLINTALOGENEYATGV	KOLLSE	1437
P70206	GLEDHPVLKEMEVOANVEKSLTLFQOLLTKKHFLITRTLEAORSFSMRDRG	VASLINTALOGENEYATGV	KOLLSE	1412
Q9UIW2	GLEDHPVLKEMEVOANVEKSLTLFQOLLTKKHFLITRTLEAORSFSMRDRG	VASLINTALOGENEYATGV	KOLLSE	1392
Q91823	-----	-----	-----	160
P51805	GLEDHPVLKEMEVOANVEKSLTLFQOLLTKKHFLITRTLEAORSFSMRDRG	VASLINTALOGENEYATGV	KOLLSE	1389
NOV12	LIEHMLSENNHKKLLRRF--TESVAEKMLTWFPTFLYKFLNECAGEPLF	MLYCAIKCOMKGPIDAI	TGEARYSLEDKI	1517
P70206	LIEHMLSENNHKKLLRRF--TESVAEKMLTWFPTFLYKFLNECAGEPLF	MLYCAIKCOMKGPIDAI	TGEARYSLEDKI	1491
Q9UIW2	LIEHMLSENNHKKLLRRF--TESVAEKMLTWFPTFLYKFLNECAGEPLF	MLYCAIKCOMKGPIDAI	TGEARYSLEDKI	1471
Q91823	-----	-----	-----	160
P51805	LIEHMLSENNHKKLLRRF--TESVAEKMLTWFPTFLYKFLNECAGEPLF	MLYCAIKCOMKGPIDAI	TGEARYSLEDKI	1468
NOV12	IRQOIDYKTLTLNCAVPEBENAPEVPVKGLCDTVOAKEKLLDAVYKGV	PYSORPKAMDMLEWROGRMARI	ILOEDV	1597
P70206	IRQOIDYKTLTLNCAVPEBENAPEVPVKGLCDTVOAKEKLLDAVYKGV	PYSORPKAMDMLEWROGRMARI	ILOEDV	1571
Q9UIW2	IRQOIDYKTLTLNCAVPEBENAPEVPVKGLCDTVOAKEKLLDAVYKGV	PYSORPKAMDMLEWROGRMARI	ILOEDV	1551
Q91823	-----	-----	-----	160
P51805	IRQOIDYKTLTLNCAVPEBENAPEVPVKGLCDTVOAKEKLLDAVYKGV	PYSORPKAMDMLEWROGRMARI	ILOEDV	1548
NOV12	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYES	MLRTASSPDSLRSRTPHITP	PDLESCTKI	1677
P70206	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYES	MLRTASSPDSLRSRTPHITP	PDLESCTKI	1651
Q9UIW2	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYES	MLRTASSPDSLRSRTPHITP	PDLESCTKI	1631
Q91823	-----	-----	-----	160
P51805	ITKIDNDWKRINTLAHYQVTDGSSVALVPKOTSAYNISNSTFTKSLRYES	MLRTASSPDSLRSRTPHITP	PDLESCTKI	1628
NOV12	NHLVNHDLHLDQREGDRCSKMVSEIYLTRLLATK--FLQKFVDDLFTET	FSTAHRGSALPLAIKYMDF	FLDEQADKHQIH	1757
P70206	NHLVNHDLHLDQREGDRCSKMVSEIYLTRLLATK--FLQKFVDDLFTET	FSTAHRGSALPLAIKYMDF	FLDEQADKHQIH	1730
Q9UIW2	NHLVNHDLHLDQREGDRCSKMVSEIYLTRLLATK--FLQKFVDDLFTET	FSTAHRGSALPLAIKYMDF	FLDEQADKHQIH	1710
Q91823	-----	-----	-----	160
P51805	NHLVNHDLHLDQREGDRCSKMVSEIYLTRLLATK--FLQKFVDDLFTET	FSTAHRGSALPLAIKYMDF	FLDEQADKHQIH	1707
NOV12	SDVRHTWKSNG--LPLRFVWVIFNPQVFDIHNSTITACLSVVAOTF	SDSCSTSEHRLGKDSPS	NLLYAKDI	1837
P70206	SDVRHTWKSNG--LPLRFVWVIFNPQVFDIHNSTITACLSVVAOTF	SDSCSTSEHRLGKDSPS	NLLYAKDI	1809
Q9UIW2	SDVRHTWKSNG--LPLRFVWVIFNPQVFDIHNSTITACLSVVAOTF	SDSCSTSEHRLGKDSPS	NLLYAKDI	1754
Q91823	-----	-----	-----	160
P51805	SDVRHTWKSNG--LPLRFVWVIFNPQVFDIHNSTITACLSVVAOTF	SDSCSTSEHRLGKDSPS	NLLYAKDI	1786
NOV12	NVERRYTADIAK--PAISDOHSEYLAPOSRLHSCNSGALHETSHIARWD	--NLLVADSKPOASRLSKLEQ		1917
P70206	NVERRYTADIAK--PAISDOHSEYLAPOSRLHSCNSGALHETSHIARWD	--NLLVADSKPOASRLSKLEQ		1886

Q9UIW2 1754
 Q91823 160
 P51805 1863
 NOV12 1925
 P70206 1894
 Q9UIW2 1754
 Q91823 160
 P51805 1871

Tables 12E-12N list the domain descriptions from DOMAIN analysis results against NOV12. This indicates that the NOV12 sequence has properties similar to those of other proteins known to contain this domain.

Table 12E. Domain Analysis of NOV12

gnl|Smart|smart00630, Sema, semaphorin domain (SEQ ID NO:113)
 CD-Length = 430 residues, 100.0% aligned
 Score = 242 bits (618), Expect = 1e-64

Query: 69 LTHLVVHEQTGEVYVGAVNRIYKLSGNLTLLRAHVTGPVEDNEKCYPPPSVQSCPHGLGS 128
 Sbjct: 1 LQNLLDDEDNGLTYVGARNRLYVLSNLISEAEVKTGPVLSSPDCEEC--VSKGKDPP-- 56
 Query: 129 TDNVNK-LLLLDYAANRLACGS-ASQGICQFLRLDDLFLKLGEPHHRKEHYLSSVQEAGS 186
 Sbjct: 57 TDCVNFIRLLLDYNADHLLVCGTNAPQPVCRILINLGNLDRLE-EVGRESGRGRCPDPQHN 115
 Query: 187 MAGVLIAGPPGQQAQKLVFGTPIID--GKSEYFPTLSSRRLMANEEDADMFGFVYQDEFVS 244
 Sbjct: 116 STAVLVGD-----ELYVGTVADFGSDPAIYRSLSVRRLLKGTSG-----PSLRTVL 161
 Query: 245 SQLKIPSDTLKFPFADIIYVYSFRSEQFVYYLTLQLDTQLTSPDAAGEHFFTSKIVRLC 304
 Sbjct: 162 YDSRWLN-----EPNFVYAFESGDFVYF----FFRETAVEDENCGKAVVSRVARVC 208
 Query: 305 VDD-----PKFYSYVEFPICG---EQAGVEYRLVQDAYLSRPGRALAHQLGLAEDED 353
 Sbjct: 209 KNDVGGPRSLSKKWTSLFKARLECSVPGEFFPYFNEAQAAFLLPAG-----SESDD 259
 Query: 354 VLFTVFAQQQKNRVKPPKESALCLFTLRAIKEKIKERIQSCYRGECKLSL----PWLLNK 409
 Sbjct: 260 VLYGVFSTS----SNPIPGSAVCAFSLSDINAVFNEPFKECETGNSQWLPYPRGLVPFPR 315
 Query: 410 ELGCINSPLQI----DDDFC-GQDFNQPLGGTVTIEGTPLFV--DKDDGLTAVAA----Y 458
 Sbjct: 316 PGTCNPNTPLSSKDLPPDVLNFIKTHPLMDEVVQPLTGRPLFVKTDNLYLLTSIAVDRVRT 375
 Query: 459 DYRGRTVVFAGTRSGRIKILVDLSNPGGRPALAYESVVAQEGSPILRDLVLSPNH 514
 Sbjct: 376 DGGNYTVLFLGTS DGRILKVVLSRSSSSSESVVLEEISVFDPGSPV-SDLVLSPPK 430

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.